

VIETNAM NATIONAL UNIVERSITY OF HOCHIMINH CITY
THE INTERNATIONAL UNIVERSITY
SCHOOL OF COMPUTER SCIENCE AND ENGINEERING



**DEEP LEARNING-BASED SKIN CANCER
DETECTION: OPTIMIZATION AND WEB-
BASED DEPLOYMENT FOR ENHANCED
DIAGNOSTIC ACCURACY**

by

NGUYEN THANH BINH

ITDSIU20056

A thesis submitted to the School of Computer Science and Engineering
in partial fulfilment of the requirements for the degree of
Bachelor of Data Science.

Ho Chi Minh City, Vietnam

2025

**Deep Learning-Based Skin Cancer Detection: Optimization and Web-Based
Deployment for Enhanced Diagnostic Accuracy.**

APPROVED BY:

Dr. Mai Hoang Bao An,

Assoc. Prof. Nguyen Thi Thuy Loan,

Dr. Nguyen Trung Ky.

THESIS COMMITTEE
(Whichever applies)

ACKNOWLEDGMENTS

This thesis marks a vital point in my educational journey, accomplished with the guidance, support, and motivation of those who have left a lasting impression on my growth.

I would like to begin by expressing my heartfelt appreciation to the faculty members at International University (IU), whose commitment to academic excellence has inspired me throughout my studies. Their thoughtful insights and willingness to share their expertise have greatly enhanced my understanding and encourage my passion for learning.

To my family, words cannot fully capture the depth of my gratitude. Their unwavering support and belief in my abilities have been my anchor through challenging times. From their patient listening to their encouragement, they have been the foundation of my strength and determination.

I owe a special debt of gratitude to my advisor, Dr. Mai Hoang Bao An, for his exceptional mentorship. His sharp insights, constructive feedback, and genuine interest in my research have been invaluable in shaping the quality of this work. I am deeply appreciated for his encouragement explore ideas beyond my comfort zone, which has greatly enriched my academic growth.

I would also like to acknowledge my friends, who have been a source of strength and companionship during this journey. Their humour and their encouragement have turned stressful days into moments of relief, and their faith in me has been a constant reminder to keep pushing forward.

Lastly, I want to extend my sincere thanks to everyone who, in ways both big and small, has contributed to my success. Whether through thoughtful discussions, moral support, or simply believing in my potential, their kindness has left a memorable mark on my journey.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	2
TABLE OF CONTENTS	3
LIST OF FIGURES	6
LIST OF EQUATIONS.....	8
LIST OF TABLES.....	9
ABSTRACT	10
CHAPTER 1	11
INTRODUCTION	11
1.1. Background.....	11
1.2. Problem Statement.....	14
1.3. Scope and Objective.	15
1.4. Assumption and Solution.....	16
1.5. Structure of Thesis.	17
CHAPTER 2	18
LITERATURE REVIEW	18
2.1. Skin cancer diagnosed with AI apps and websites.	18
2.1.1. Skin Vision.....	18
2.1.1.1. Key Features of Skin Vision.....	19
2.1.1.2. Advantages of Skin Vision.	22
2.1.1.3. Disadvantages of Skin Vision.....	22
2.1.2. MiiSkin.....	22
2.1.2.1. Key Features of MiiSkin.....	23
2.1.2.2. Advantages of MiiSkin.	25
2.1.2.3. Disadvantages of MiiSkin.....	25
2.1.3. Molescope.	25
2.1.3.1. Key Features of MoleScope.....	26
2.1.3.2. Advantages of MoleScope.	28
2.1.3.3. Disadvantages of MoleScope.....	28
2.2. Related work.....	29
2.3. Convolutional Neural Network (CNN) in Machine Learning.....	31

2.3.1.	Overview of CNN architecture	32
2.3.1.1.	Input Layer.....	32
2.3.1.2.	Convolutional Layer	32
2.3.1.3.	Pooling Layer.....	33
2.3.1.4.	Normalization Layer.....	33
2.3.1.5.	Flattening Layer.....	34
2.3.1.6.	Fully Connected Layer.....	35
2.3.1.7.	Activation Function.	35
2.3.2.	DenseNet architecture.....	36
2.3.2.1.	Key features of the DenseNet model	37
2.3.3.	Inception architecture.....	39
2.3.3.1.	Key features of the Inception model.....	40
2.4.	Techniques for website implementation.....	41
2.5.	Techniques for machine learning models and API service.	42
CHAPTER 3		43
METHODOLOGY		43
3.1.	Proposed method for skin lesions prediction.....	43
3.2.	Dataset.....	44
3.3.	Data visualisation.....	45
3.4.	Data pre-processing and Data augmentation.	48
3.4.1.	Data Pre-processing.....	49
3.4.2.	Data Augmentation techniques.....	49
3.5.	Deep learning skin lesion predictions models.	52
3.5.1.	DenseNet201 model.	53
3.5.2.	InceptionV3 model.	53
3.5.3.	Xception model.....	53
3.5.4.	The Vision Transformer model.	54
3.5.5.	Performance metrics.....	54
CHAPTER 4		56
IMPLEMENT AND RESULTS		56
4.1.	Experimental setup.	56
4.2.	Deep Learning models results.	60
4.2.1.	Densenet201 model results.....	60
4.2.2.	InceptionV3 model results.	63

4.2.3.	Xception model results.....	66
4.2.4.	ViT model results.....	69
4.3.	Website Implementation.....	72
4.3.1.	Home page.....	73
4.3.2.	Information page.....	75
4.3.3.	Prediction page.....	80
CHAPTER 5		83
DISCUSSION AND EVALUATION		83
5.1.	Discussion.....	83
5.2.	Comparison.....	84
5.3.	Evaluation	84
CHAPTER 6		86
CONCLUSION AND FUTURE WORK		86
6.1.	Conclusion	86
6.1.	Future work.....	86
REFERENCES		88
APPENDIX A.....		93
APPENDIX B		94

LIST OF FIGURES

Figure 1.1:The example of Melanoma diseases.	11
Figure 1.2:The example of Basal cell carcinoma (BCC) diseases.	12
Figure 1.3: The example of skin cancer damaged people's appearance.	13
Figure 1.4: Closer look of how skin cancer lesions damaged a person's visual appearance....	13
Figure 1.5: The structure of thesis.....	17
Figure 2.1:The SkinVision logo.	18
Figure 2.2: The SkinVision user interface design.	19
Figure 2.3: The process to get the diagnosed in SkinVision.	20
Figure 2.4:The result after users scanned their skin and uploaded images in SkinVision.	20
Figure 2.5: The report of their skin condition in SkinVison.	21
Figure 2.6: The warnings and advices user interface design in SkinVision.....	21
Figure 2.7: The Miiskin logo.....	22
Figure 2.8:The Miiskin user interface design.....	23
Figure 2.9:The skin analysis procedure in MiiSkin.....	23
Figure 2.10:The tracking changes in lesion over time in MiiSkin.	24
Figure 2.11: The consulting dermatologist in the local area function in MiiSkin.....	24
Figure 2.12:The MoleScope logo.	26
Figure 2.13:The MoleScope user interface design.	26
Figure 2.14:The MoleScope camera in MoleScope.	27
Figure 2.15:An example of how cameras work in MoleScope.	27
Figure 2.16:An example of skin lesion analysis techniques in MoleScope.	28
Figure 2.17:The overview of CNN architecture [29].	31
Figure 2.18:The comparison table of all DenseNet architectures [31].	37
Figure 2.19 : The Inception overall architecture [35].....	40
Figure 3.1:Proposed methodology in this study.	43
Figure 3.2:The examples of sample lesions in the dataset [55].....	45
Figure 3.3: The bar chart illustrates how skin lesion types are distributed within the dataset.	45
Figure 3.4:The bar chart shows the distribution of diagnostic confirmation methods for diseases.	46
Figure 3.5:The bar chart displays the counts of the body areas that have skin lesions.	47
Figure 3.6:The bar chart presents the distribution of all skin lesions types	47
Figure 3.7:The frequency chart above presents the age distribution of all patients.	48
Figure 3.8:Examples of data augmentation techniques on the skin lesions dataset.	50
Figure 3.9:The code of data augmentation technique code.	51
Figure 3.10:The code for ViT augmentation.	52
Figure 4.1:The code for three deep learning models.	59
Figure 4.2:The code for ViT models.	60
Figure 4.3:The confusion matrix for the DenseNet201 model.....	60
Figure 4.4:Performance evaluation of DenseNet201 model.	60
Figure 4.5:The classification report on DenseNet201.	61
Figure 4.6:The confusion matrix for the InceptionV3.....	63
Figure 4.7:Performance evaluation of InceptionV3 model.	63

Figure 4.8: The classification report on the Inception V3 model.....	64
Figure 4.9: The confusion matrix for the Xception model.....	66
Figure 4.10: Performance evaluation of Xception.....	66
Figure 4.11: The classification report on the Xception model.....	67
Figure 4.12: The confusion matrix for the ViT model.....	69
Figure 4.13: Performance evaluation of ViT.....	70
Figure 4.14: The classification report on the ViT model.....	70
Figure 4.15: The website's logo.....	73
Figure 4.16: The website's landing page user interface design.....	75
Figure 4.17: The header of the information page.....	75
Figure 4.18: The general knowledge section.....	76
Figure 4.19: The Fitzpatrick skin classification system section.....	77
Figure 4.20: The ABCDE rule section.....	77
Figure 4.21: The additional information section.....	78
Figure 4.22: The additional information section.....	78
Figure 4.23: The skin disease gallery section.....	79
Figure 4.24: The message for the users to take care of their skin health.....	79
Figure 4.25: The upload picture user interface.....	80
Figure 4.26: The error message user interface design.....	81
Figure 4.27: The prediction result user interface design.....	82
Figure 4.28: The references link to Wikipedia when the users want to know further information.....	82

LIST OF EQUATIONS

Equation 2.1:The formula for convolutional layers [21].....	32
Equation 2.2:The formula for maximum pooling [22].....	33
Equation 2.3:The formula for average pooling [23].....	33
Equation 2.4:The formula batch normalization [25]	34
Equation 2.5:The formula for a fully connected layer [22].....	35
Equation 2.6:The formula for ReLU [31].....	36
Equation 2.7:The forumla for Softmax [21]	36
Equation 2.8:The formula Dense Block [31].....	37
Equation 2.9:The formula for Transition layer [27].....	38
Equation 2.10:The formula for Growth Rate [32].....	39

LIST OF TABLES

Table 3.1:The setup parameters for three deep learning models trained in the thesis.....	57
Table 5.1:Comparison with other models.....	85

ABSTRACT

Addressing complex challenges across diverse fields has been revolutionized by the growing adoption of machine learning techniques. With rapid advancements in technology, its applications in image recognition have seen significant progress. However, achieving high-performance machine learning models while ensuring smooth integration into user-friendly web platforms remains a significant challenge that requires innovative solutions.

This thesis aims to advance deep learning models specifically designed for the diagnosis of skin cancer. A unique neural network architecture is proposed and its performance carefully evaluated against four well-established frameworks: Xception, InceptionV3, DenseNet201, and ViT model. The most effective model is subsequently integrated into a web-based platform, providing an accessible and intuitive solution for end-users. The research outlines a structured approach consisting of three main steps:

- **Dataset Collection:** A powerful and diverse dataset is a cornerstone for building effective machine learning model. For this study, the HAM10000 dataset from Kaggle serves as the primary data source. This dataset contains an extensive collection of labeled skin cancer images, providing a reliable foundation for training and validation. (Dataset link: <https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000> [55])
- **Data Pre-processing:** This stage involves preparing the dataset for analysis by cleaning irrelevant attributes, eliminating redundant data, and transforming images into a compatible structure for deep learning models. Additionally, techniques are applied to address class imbalance issues effectively.
- **Model Deployment and Continuous Improvement:** The final phase involves deploying the optimized classification model on a web-based application. Feedback process are incorporated to continuously refine both the model and the user interface, ensuring accurate predictions and a smooth experience for users. The continuous process enables the system to adapt and improve over time based on real-world interactions.

CHAPTER 1

INTRODUCTION

1.1. Background.

Cancerous skin lesions have become a growing international public health concern in recent years, driven by the increasing frequency and substantial consequences associated with these conditions. In the United States alone, it ranks as one of the most frequently occurring cancers, with approximately 5.4 million new cases reported annually [1]. Of the different types of skin cancer, melanoma is particularly concerning due to its aggressive nature and high mortality rates, owing to its rapid progression and high mortality rates. In 2022 alone, melanoma was responsible for an estimated 97,000 new cases in the United States [2]. On a global scale, the impact of skin cancer is equally alarming, with over 60,000 deaths and 132,000 new diagnoses reported each year. This highlights the significant burden it places on healthcare systems worldwide [3].

Melanoma, recognized as the deadliest form of skin cancer, poses critical risks if not detected and treated early. When left untreated, it can move to essential organs such as the lungs, liver, brain, or lymph nodes, leading to organ failure and often fatal outcomes [4]. Treatment for advanced melanoma remains challenging, with limited success rates and substantial financial costs, making it a significant concern for both patients and healthcare providers. Beyond the physical implications, melanoma negatively affects patients' life expectancy, includes emotional and psychological stress, and creates economic challenges for affected families [5].



Figure 1.1: The example of Melanoma diseases.



Figure 1.2:The example of Basal cell carcinoma (BCC) diseases.

In Vietnam, while malignant skin conditions are relatively uncommon compared to other diseases, it remains a pressing public health concern. According to the World Health Organization (WHO), Vietnam sees around 165,000 new cancer diagnoses and approximately 115,000 cancer-related deaths annually [6]. While specific national data on skin cancer is limited, global figures indicate that in 2020, Vietnam recorded 1,230 new cases of skin cancer, representing roughly 1 case out of every 100,000 individuals [7]. The country's tropical climate, which is characterized by consistently high levels of ultraviolet (UV) radiation, particularly during the peak seasons, significantly increases the likelihood of skin cancer development. These environmental factors highlight the need for preventive strategies, such as public education on sun protection, proper clothing, and regular skin checks for early detection of any irregularities.

Apart from melanoma, which is regarded as the most aggressive type of skin tumors, other forms such as basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) also contribute to the burden of the disease though BCC and SCC are generally less aggressive than melanoma, they can still cause lasting damage if left untreated, especially when they affect sensitive areas like the face [8]. These conditions greatly affect the overall well-being of individuals, often leading to a decline in self-esteem and increased anxiety due to the visible nature of skin lesions. This is particularly true for those who experience tissue loss in visible areas, leading to feelings of embarrassment in social interactions. The psychological effect associated with skin cancer can result in social isolation, anxiety, and, in some cases, a decline in overall health [9]. Furthermore, skin cancer lesions not only cause cosmetic issues but can

also lead to infections, pain, and reduced movement, especially when located on particularly vulnerable regions like the arms, legs, or neck [10].



Figure 1.3: The example of skin cancer damaged people's appearance.



Figure 1.4: Closer look of how skin cancer lesions damaged a person's visual appearance.

The development of skin cancer is strongly associated with sustained exposure to ultraviolet radiation, either through natural means such as sunlight or through artificial methods like tanning equipment, is a primary contributor to skin cancer incidence. According to Cancer Research UK, both natural sunlight and artificial UV sources, such as tanning beds, contribute significantly to the increasing rates of skin cancer [11]. This highlights the urgent need for preventive strategies, including public education on sun safety, the use of protective measures such as sunscreen and clothing, and reducing the amount of time spent under UV light.

The field of skin cancer detection has experienced a remarkable shift due to advancements in artificial intelligence and deep learning by introducing innovative and effective diagnostic solutions. These technologies have enabled the development of diagnostic tools that are not only faster but also more reliable and accurate. AI-driven systems, in particular, demonstrate exceptional abilities in analysing large datasets of dermoscopy images, allowing for the rapid and precise identification of skin abnormalities compared to traditional methods.

The integration of deep learning algorithms into diagnostic platforms offers a ground-breaking approach to addressing challenges in early detection. Such platforms have the potential to empower individuals to monitor their skin health independently while providing valuable recommendations for skin care and preventive measures. By utilizing extensive datasets of dermoscopy images, these systems significantly enhance detection accuracy, making skin cancer screening more accessible to the general public. This innovation creates a connection between traditional diagnostic approaches and modern technological advancements, ultimately improving patient outcomes and promoting early intervention [12].

In conclusion, SkinGuard applies artificial intelligence to support early skin cancer detection, enabling timely measures to protect skin health and improve overall well-being.

1.2. Problem Statement.

In the past, diagnosing skin cancer through traditional methods posed significant challenges for patients and healthcare providers. These challenges include:

- **Limited access to dermatologists:** Patients in rural or unpopular areas often struggle to access specialized care about skin conditions. Long waiting times for appointments and diagnoses further delay treatment, leading to poorer health outcomes.
- **Time-consuming procedures:** Traditional diagnostic methods, such as biopsies, require multiple visits for testing, consultations, and follow-ups. This process not only takes time but also increases patients' emotional and physical stress.
- **High financial costs:** Procedures like biopsies are costly, particularly for patients without health insurance. When combined with treatment expenses, these costs create a substantial financial burden.
- **Inconsistent diagnostic accuracy:** Even skilled dermatologists encounter difficulties in distinguishing benign lesions from malignant ones because of their similar visual characteristics. This can result in miss diagnoses or inappropriate treatment plans.

Advances in deep learning and computer technology have transformed the methods used for diagnosing and identifying skin cancer. These tools address many of the limitations of traditional methods, offering potential for improved accuracy, accessibility, and efficiency. However, current systems still face several challenges:

- **Limited diagnostic scope:** Many AI models are designed to detect specific skin conditions, limiting their ability to provide a comprehensive diagnosis across a wide range of dermatological issues.
- **Inconsistent accuracy:** Even experienced dermatologists often face challenges in distinguishing benign from malignant skin lesions due to the similar visual appearance of these conditions. This challenge can result in misdiagnosis or ineffective treatment plans, potentially affecting patient care and outcomes.
- **User interface challenges:** Poorly designed interfaces can make it difficult for patients to use AI diagnostic tools effectively, reducing their accessibility and overall usability.
- **Data quality and availability:** Developing AI models requires diverse and high-quality datasets. However, privacy laws and ethical considerations often restrict access to sufficient medical data, model training and improvement.
- **Privacy and security concerns:** The increasing reliance on online diagnostic platforms raises concerns about the security of sensitive patient data, such as medical histories and skin images.

Building a deep learning-based website for skin cancer detection can address many of the issues faced by patients and dermatologists. While current AI systems have limitations, ongoing advancements are continuously improving their accuracy, accessibility, and security. These developments create a promising outlook for AI-driven healthcare solutions, assuring earlier detection and better management of skin cancer.

1.3. Scope and Objective.

The proposed skin cancer detection website is designed to serve a diverse audience, including individuals seeking early diagnosis, healthcare professionals, and dermatologists involved in initial assessment stages. It suitable for users of all ages, from patients requiring first-time evaluations to private clinics and dermatology hospitals. Additionally, the platform aims to promote public awareness of skin cancer prevention through educational resources, practical skincare tips, and actionable advice for maintaining skin health.

The website incorporates a sophisticated deep learning model trained on an extensive dataset. This model delivers highly accurate diagnostic results, empowering users to take

appropriate actions for protecting their skin. Alongside the diagnostic tool, the platform provides personalized skincare advice tailored to the user's skin type and condition.

Furthermore, the website offers essential knowledge about skin lesions, including early warning signs of skin cancer, helping users understand their conditions better and take proactive measures. By combining cutting-edge AI technology with user-focused features, the website reduces the gap between traditional diagnostic methods and innovative healthcare solutions, creating a reliable, secure, and accessible tool for patients and medical professionals.

In summary, building a skin cancer detection website requires combining effective deep learning models with simple and user-friendly features. The main goals are to offer accurate diagnostic tools and increase public knowledge about preventing and managing skin cancer.

1.4. Assumption and Solution.

Designing an application for skin cancer detection involves addressing various challenges and needs throughout the diagnostic process. Below are key assumptions and specialized solutions aimed at improving user experience and diagnostic efficiency:

- **Assumption:** Users lack knowledge about skin lesions and cancer types.

Solution: Create a user-friendly educational section that provides detailed insights into skin cancer, including common lesion types and early symptoms. This section will feature visual aids to enhance understanding and encourage proactive skin health monitoring.

- **Assumption:** Concerns regarding the accuracy of AI diagnostic tools.

Solution: Employ advanced AI and deep learning architectures, supported by a dataset containing over 10,000 images, to ensure reliable and precise diagnostic results.

- **Assumption:** Users expect quick diagnostic results response.

Solution: Optimize the system by training multiple deep learning models and selecting the one that achieves the best balance between accuracy and computation time, enabling faster predictions.

- **Assumption:** Users may struggle to understand complex medical results.

Solution: Simplify diagnostic results by using clear, non-technical language and visual aids, helping users understand their condition without confusion.

- **Assumption:** Users upload low-quality or incorrect images.

Solution: Include step-by-step guidance on capturing high-quality images and provide a retry option for users to upload other images if necessary.

- **Assumption:** Users require follow-up care or recommendations.

Solution: After providing diagnostic results, the system will include suggestions for further actions, such as recognizing additional symptoms or consulting local clinics for a more thorough examination.

The assumptions and solutions outlined must undergo continuous evaluation and refinement to ensure the application becomes a reliable and efficient tool for dermatologists and individuals seeking early diagnosis of skin conditions.

1.5. Structure of Thesis.

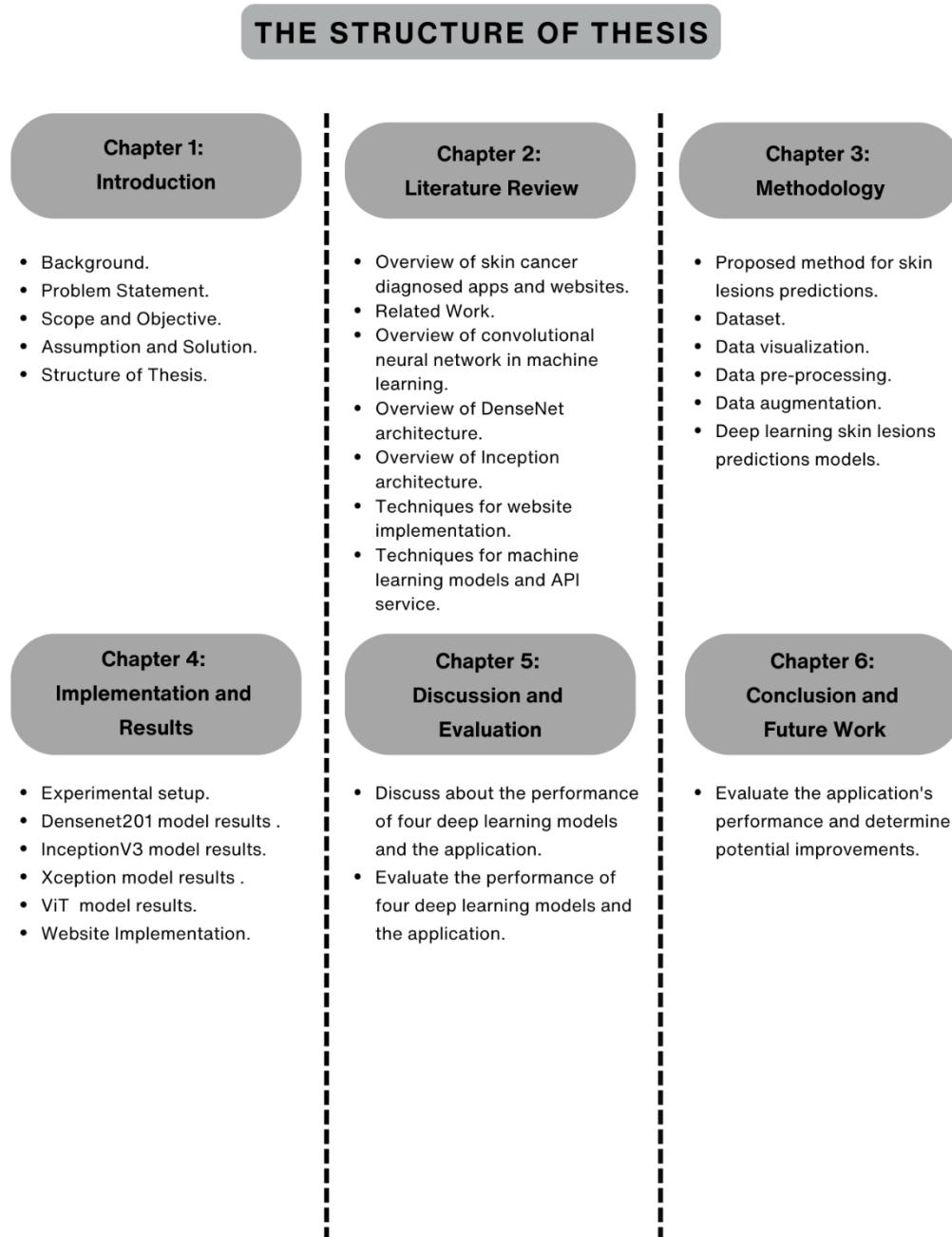


Figure 1.5: The structure of thesis.

CHAPTER 2

LITERATURE REVIEW

2.1. Skin cancer diagnosed with AI apps and websites.

In today's era of rapidly advancing digital technology, artificial intelligence has brought significant progress to the medical field. One important application of AI is assisting in the diagnosis of skin cancer, a condition with a high occurrence rate that can be treated effectively if detected early. AI-powered web tools and mobile applications enable users to check their skin condition at home, reducing the need for hospital visits. These tools analyse skin images using AI to identify potential signs of skin cancer and provide early warnings.

2.1.1. Skin Vision.

SkinVision is a mobile application that utilizes artificial intelligence to detect skin cancer from photos. This app enables users to self-check their skin for potential skin cancer risks without the need to visit a medical facility. By capturing images of moles, acne marks, or skin lesions, the app analyses these photos to assess the likelihood of skin cancer.

SkinVision was created with the goal of helping users identify early signs of skin cancer, which can lead to seek medical attention when necessary. The application employs advanced machine learning algorithms, which have been validated by international medical organizations, to ensure the accuracy of its analysis. After processing the images, the app categorizes the risk level of the skin condition, guiding users on when to seek professional medical advice [13].



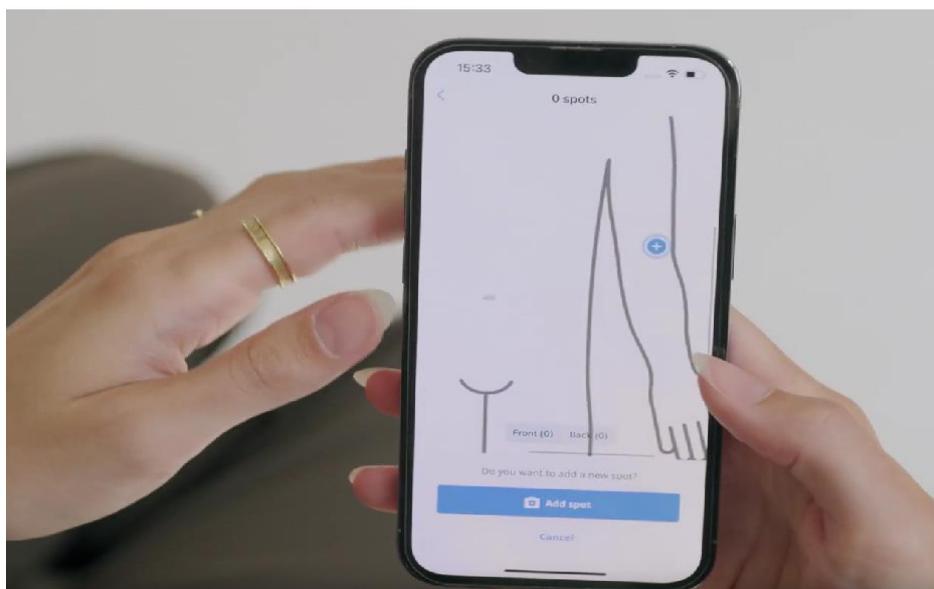
Figure 2.1: The SkinVision logo.



Figure 2.2: The SkinVision user interface design.

2.1.1.1. Key Features of Skin Vision.

- **Scan and analyse images:** Users can upload photos of acne spots, moles, and other skin changes to the app. SkinVision uses AI technology to assess the properties of these photos and predict the signs of skin cancer.



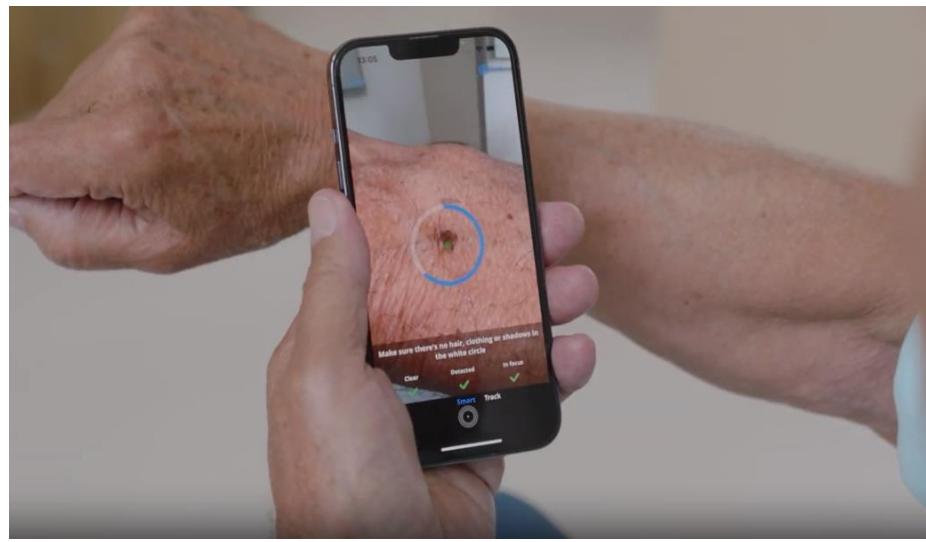


Figure 2.3: The process to get the diagnosed in SkinVision.

- **Risk assessment:** After analysing the image, the app will provide an assessment of the risk level of that pimple or mole, from "no concern" to "high risk". This helps users identify danger signs and user can decide whether to see a doctor or not.

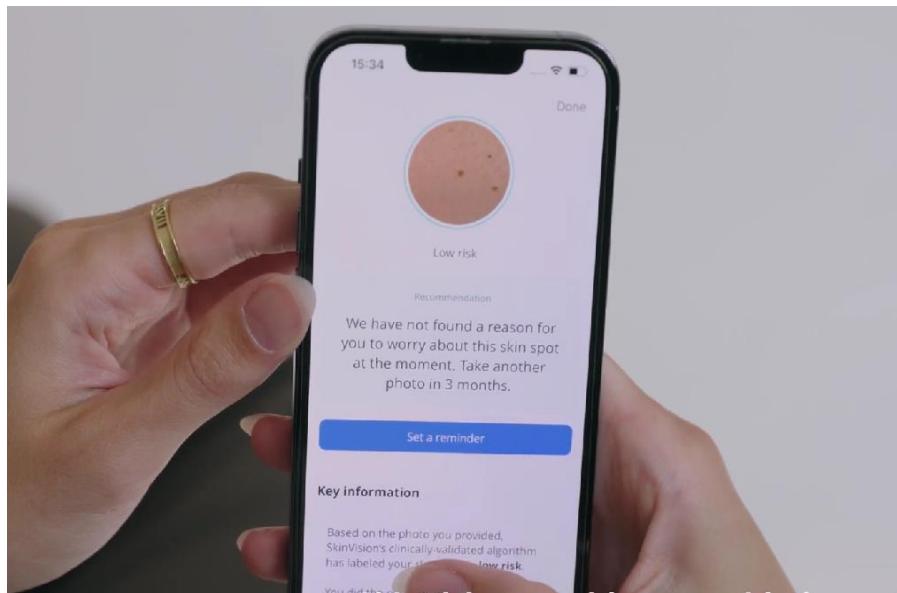


Figure 2.4: The result after users scanned their skin and uploaded images in SkinVision.

- **Tracking history:** SkinVision allows users to save previously obtained photos and follow mole changes over time. This is especially beneficial for spotting unnormal changes in acne or moles, which are a key indicator of skin cancer.

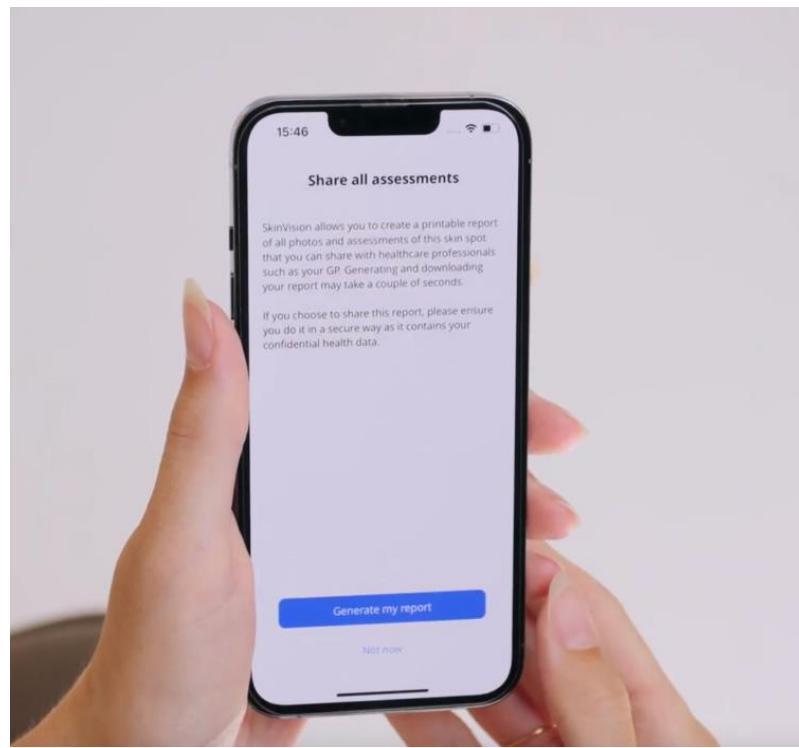


Figure 2.5: The report of their skin condition in SkinVision.

- **Warnings and advice:** If serious acne spots are found, the system will alert users and recommend that they visit a doctor or dermatologist. Furthermore, the app provides information on dermatologists around you, making it easy to locate professional help.

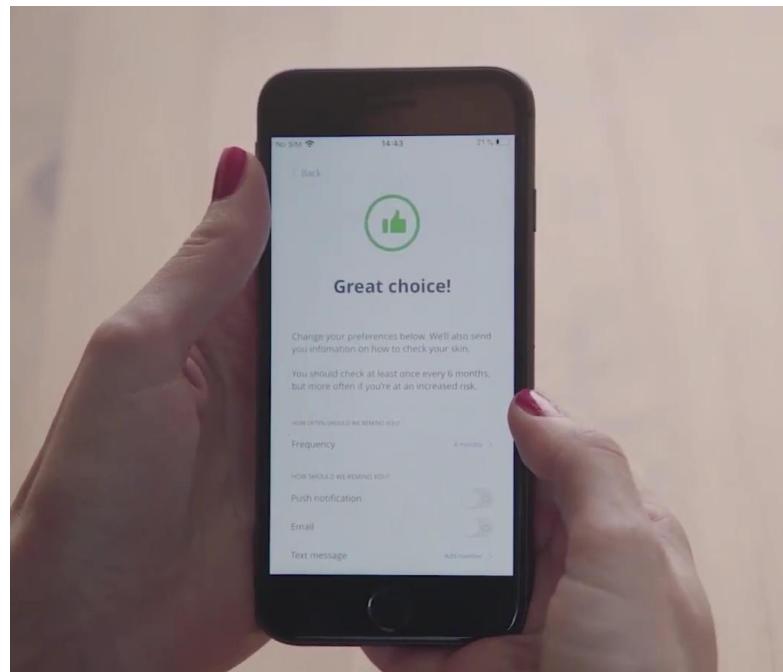


Figure 2.6: The warnings and advices user interface design in SkinVision.

2.1.1.2. Advantages of Skin Vision.

- SkinVision has been verified and strictly validated to ensure that it can reliably detect indicators of skin cancer.
- SkinVision customers can share analysis results as well as images of lesions with dermatologists and other medical specialists.
- SkinVision generates a complete analysis of the user's skin damage, including recommendations on how to continue monitoring and when to see a doctor.

2.1.1.3. Disadvantages of Skin Vision.

- To use the app properly, users must have a mobile device with a high-quality camera capable of producing clear, detailed pictures that are not covered by anything. If the device is not capable of producing high-quality photos, the program may not function properly, and the analysis findings may be inaccurate.
- SkinVision's primary function is melanoma detection, but it may struggle with certain other types of skin abnormalities.
- SkinVision charges users for the diagnosis of skin lesions and thorough results reporting services. Although there is a free version, the complete functionality requires users to pay, which can be a barrier for some individuals.

2.1.2. Miiskin.

Miiskin is a mobile app and web platform designed to help users monitor their skin for signs of skin cancer, particularly melanoma. The app allows users to track skin changes by analysing photos of moles and other skin lesions. Miiskin's primary purpose is to offer a simple solution for detecting early-stage skin cancer indicators, encouraging users to seek professional diagnosis and treatment if necessary. By providing a user-friendly platform, Miiskin helps users stay proactive about their skin health [14].



Figure 2.7: The Miiskin logo.

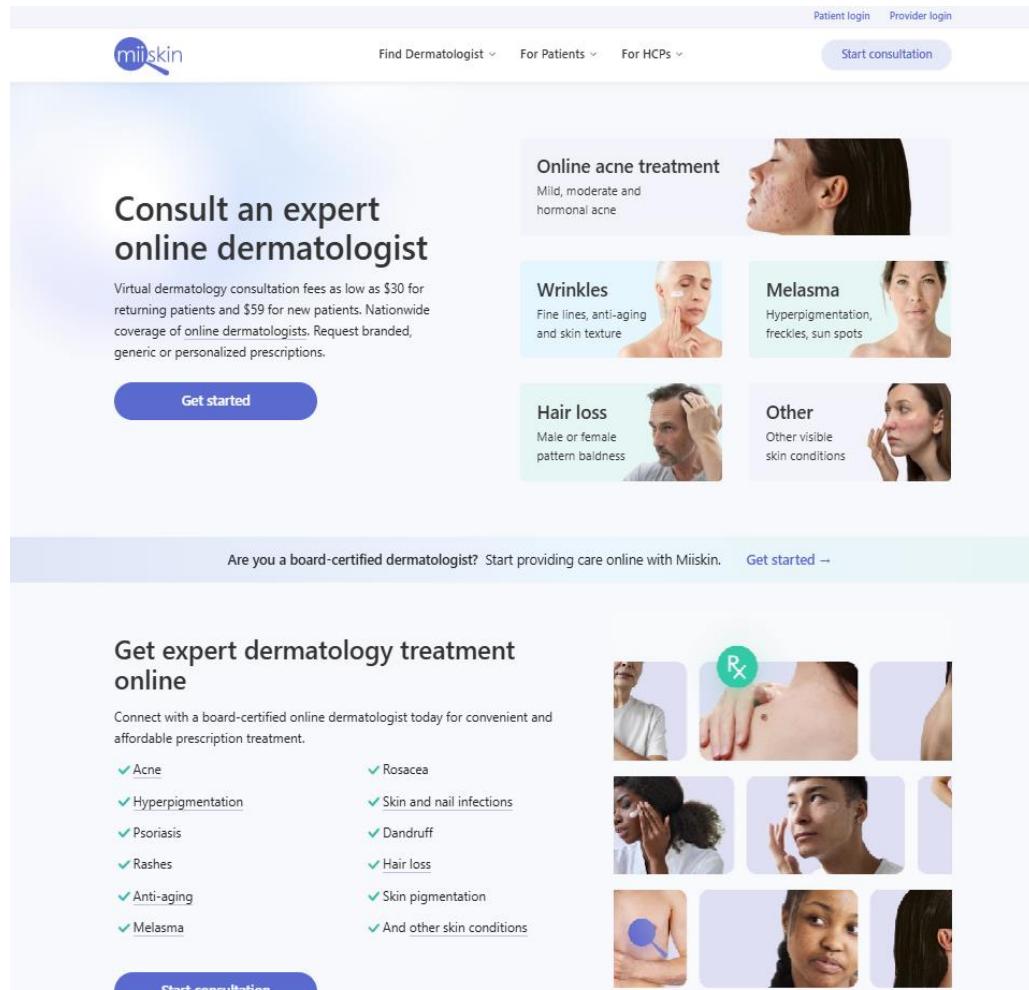


Figure 2.8: The Miiskin user interface design.

2.1.2.1. Key Features of MiiSkin.

- **Image scanning and analysis:** Miiskin lets users takes pictures of skin lesions or moles and then submit them to the app for examination. The AI algorithm assesses risk by analyzing lesions' size, shape, color, and changes over time.



Figure 2.9: The skin analysis procedure in MiiSkin.

- **Track changes in lesions over time:** Miiskin has abilities for tracking and comparing changes in moles and lesions over time. This assists users in detecting odd changes, such as changes in size or color, which are a key indicator of skin cancer.

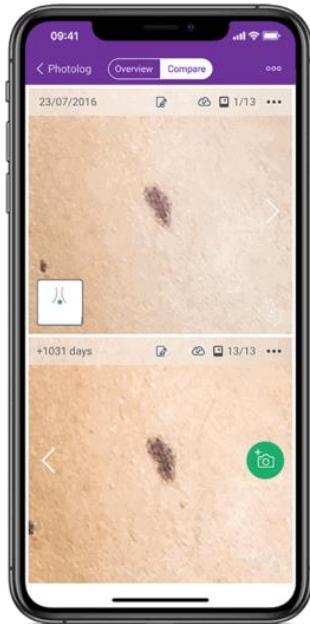


Figure 2.10: The tracking changes in lesion over time in MiiSkin.

- **Warnings and professional recommendations:** If the system detects any symptoms of danger, Miiskin will notify users and encourage them to consult a doctor or dermatologist for additional inspection. Miiskin also offers valuable tools and advice on skin care and recognizing the characteristics of skin cancer.

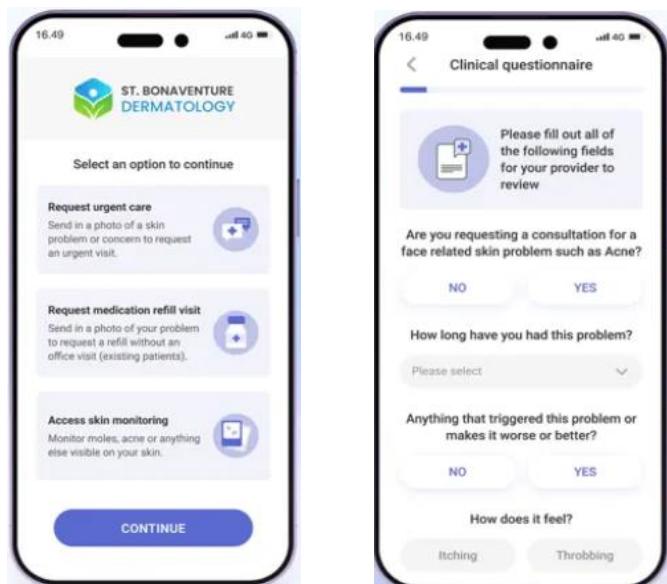


Figure 2.11: The consulting dermatologist in the local area function in MiiSkin.

2.1.2.2. Advantages of MiiSkin.

- Miiskin enables users to track changes in moles and skin lesions over time. Users can easily and quickly spot abnormal signals by taking images at regular time and comparing changes in the size, shape, and color of lesions.
- Miiskin's interface is designed for simplicity and ease of use. Users can acquire analysis findings in a matter of seconds by just taking a photo of their moles or skin lesions and uploading it to the app.
- Miiskin analyses the photos and generates a detailed report on the danger of skin lesions, ranging from "low risk" to "high risk". If abnormal signals are found, the application will recommend that users consult a dermatologist for additional evaluation. This allows users to respond quickly and seek medical attention as necessary.

2.1.2.3. Disadvantages of MiiSkin.

- To effectively analyse skin lesions with Miiskin, users must acquire clean, high-resolution photos. If the photos are unclear or not precise enough, the analysis results may be incorrect, resulting in false alarms or overlooking dangerous details.
- While Miiskin prioritizes detecting melanoma, which is the most life-threatening form of skin lesions, other categories of skin cancer prove more challenging to diagnose.
- The Miiskin software requires users to monitor their skin on a daily basis and upload images to track changes.
- Miiskin may have trouble identifying significant or very minor abnormalities. To be more specific, if a pimple's shape or color changes slightly, Miiskin may not detect or alert the user to it.

2.1.3. Molescope.

MoleScope is a mobile device and app designed to assist users in monitoring and detecting changes in their skin, including moles, acne, and other potential skin cancer-related lesions. This application is especially helpful for people who have a family background of skin cancer or a large number of moles, allowing them to monitor their skin more effectively. Moreover, users can take high-quality photos of their skin and track how these features change over time. MoleScope also works in combination with the mobile app, which provides a comprehensive view of skin conditions and alerts users when irregularities are detected. This combination of features helps users stay informed about their skin health and take appropriate actions if needed [15].



Figure 2.12: The MoleScope logo.

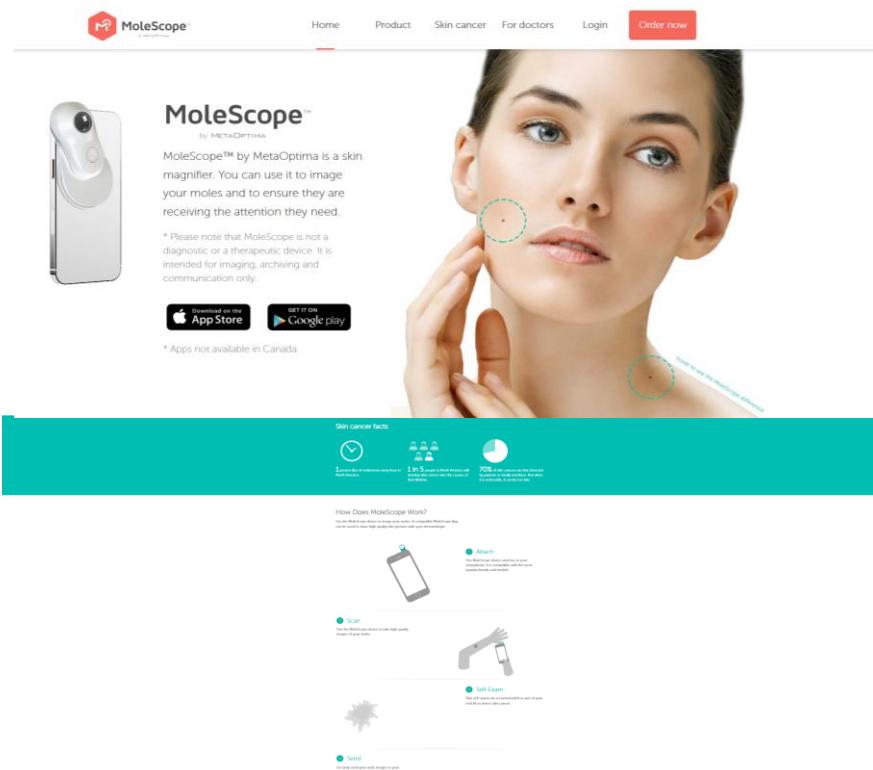


Figure 2.13: The MoleScope user interface design.

2.1.3.1. Key Features of MoleScope.

- **Special equipment included:** MoleScope Skin Dedicated Camera includes a dedicated camera that allows users to capture clear and detailed images of pimples, moles, and skin lesions. This camera's high resolution and zooming features assist in clarifying the characteristics of skin lesions, which is critical in detecting minor changes that are difficult to see with the eyes.



Figure 2.14: The MoleScope camera in MoleScope.

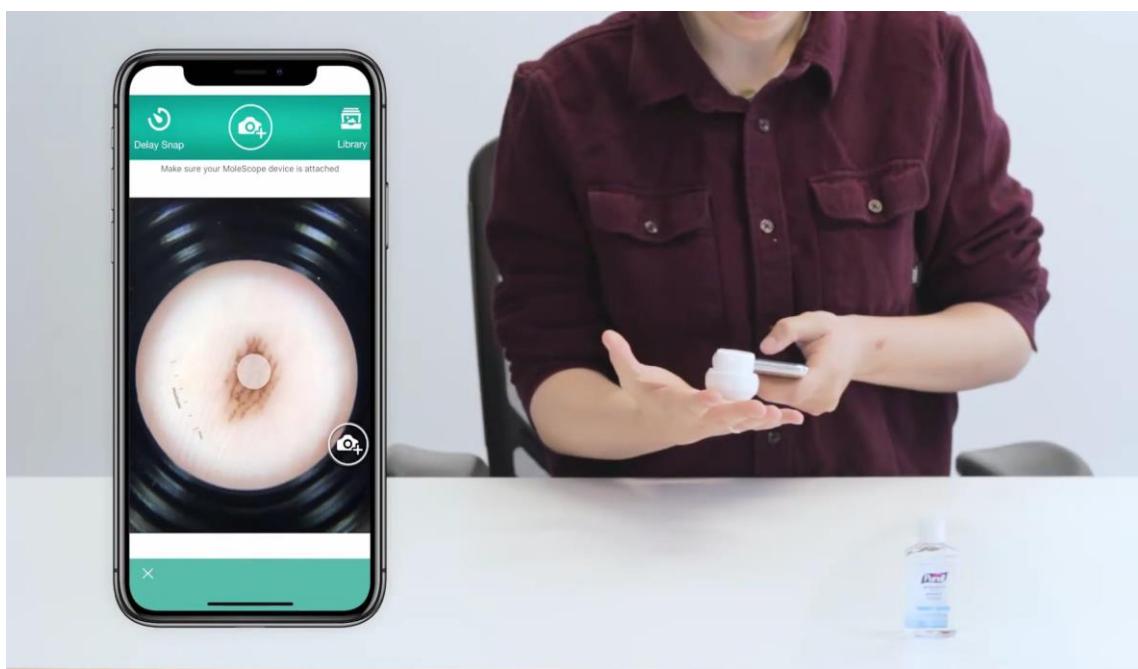


Figure 2.15: An example of how cameras work in MoleScope.

- **Skin lesion prediction:** MoleScope uses AI-powered analytical technology in its operations to scan images of acne or moles and compare them to a database of other skin lesions. AI detects unexpected signals, such as changes in the size, shape, or color of acne patches, allowing users to identify potential concerns.

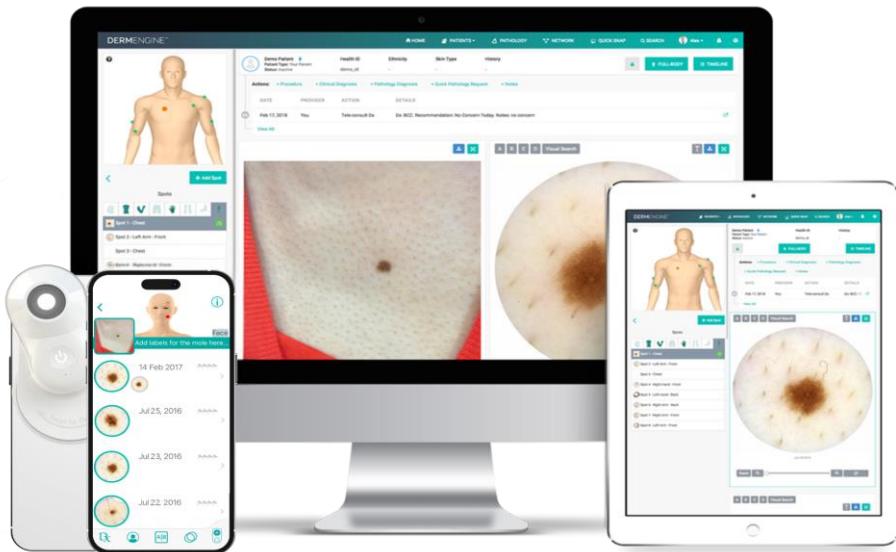


Figure 2.16: An example of skin lesion analysis techniques in MoleScope.

- **Sharing information with doctors:** MoleScope enables users to share photographs and reports with dermatologists, allowing them to receive a more thorough consultation and diagnosis. Users can transfer information about skin irregularities for doctors to evaluate remotely, which is very useful for those who live in distant places or do not have access to many dermatologists.

2.1.3.2. Advantages of MoleScope.

- MoleScope tracks changes in moles or acne to assist identify early indicators of skin cancer, particularly melanoma. Early detection can lower the risk of death and improve the chances of a successful course of therapy.
- Using a specialized camera with high resolution and zoom abilities. Therefore, MoleScope enables users to capture clear, detailed images of skin lesions with ease.
- Professional assistance by sharing data with dermatologists. Therefore, MoleScope enables customers to get professional diagnosis and advice without having to visit the hospital.

2.1.3.3. Disadvantages of MoleScope.

- The system primarily relies on the use of a dedicated camera. To ensure optimal image quality with MoleScope, users are required to utilize a specialized camera, which may be costly. This requirement could pose a challenge for individuals unwilling or unable to invest in additional equipment.

- Although MoleScope offers enormous potential for identifying skin cancer, the technology has not been thoroughly confirmed in clinical validation or formal medical studies.
- MoleScope is mainly designed to detect skin lesions in individuals with lighter skin tones. It may not provide reliable results for those with darker or uneven skin tones.

2.2. Related work

Deep learning has shown considerable promise in supporting dermatologists with skin lesion analysis, leading to faster and more precise diagnoses. Considering the widespread of skin cancer globally, early identification plays an indispensable role in optimizing treatment strategies. The application of CNNs, particularly when combined with extensive datasets like HAM10000 demonstrates their proficiency in detecting detailed patterns and accurately classifying diverse skin lesions, thereby contributing significantly to early diagnosis methods. However, despite these advancements, challenges remain in addressing rare lesion types, improving generalization, and classifying hidden features in complex cases. Below is an overview of significant studies that address both the successes and challenges in this field.

Akter et al. (2023) conducted a comprehensive evaluation of a deep learning framework designed for the precise classification of skin lesions, utilizing the HAM10000 medical dataset, which contains 10,015 high-resolution dermatoscopy images across seven lesion categories. Among the models evaluated, ResNet50 showed the best performance with an accuracy of 82%, due to its residual connections that improved training and optimization. MobileNetV2 followed with an accuracy of 87%, offering a lightweight and efficient solution suitable for devices with limited computing resources. Xception achieved an accuracy of 88%, using depthwise separable convolutions to balance performance and computational demands. In contrast, VGG16, an older CNN architecture, obtained the lowest accuracy at 77%, likely due to its simpler design and a tendency to overfit on this complex dataset. ResNet50 also performed strongly in additional metrics, such as a weighted F1-score of 0.81, precision of 0.80, and recall of 0.82. However, its large size and higher computational requirements may limit its use in real-time applications. MobileNetV2, with fewer parameters and lower energy needs, offers a practical solution for real-world deployment. This study highlights the importance of balancing accuracy and efficiency when selecting models and encourages further research into optimization techniques, such as model simplification and data reduction, to improve the usability of these models in medical imaging [16].

Yashwant Ingle and Nuzhat Faiz Shaikh (2024) conducted a comprehensive analysis comparing the performance of traditional CNNs with the VGG16 architecture for identifying skin cancer, using data from the HAM10000 collection. The evaluation of these models relied on key metrics, including accuracy, sensitivity, and specificity. VGG16, known for its deeper layers and ability to extract detailed features, achieved an accuracy of 89%, which marked a substantial improvement over the 67% accuracy of the traditional CNN. However, despite this relative advantage, VGG16 underperformed when evaluated against highly sophisticated models such as DenseNet and ResNet. One significant limitation was its reduced effectiveness in detecting rare lesion types, largely due to class imbalance within the dataset. This study highlighted the critical role of data augmentation techniques, including rotations, scaling, and image blurring, to improve model performance and address these imbalances. While VGG16 demonstrated superior results compared to traditional CNNs for this dataset, its computational demands posed challenges in resource-constrained environments. Ultimately, the researchers concluded that although VGG16 is a suitable option for moderate datasets like HAM10000, its performance could be further enhanced through advanced techniques such as hyperparameter optimization, improved data balancing, and the application of transfer learning [17].

Chaturvedi et al. (2019) studied the application of MobileNet, a lightweight convolutional neural network, for classifying skin lesions with the HAM10000 dataset, which includes 10,015 dermoscopic images distributed across seven different categories. The model, initially trained on ImageNet, was fine-tuned for this specific classification task, with pre-processing steps like resizing and data augmentation implemented to improve performance and address class imbalance. MobileNet achieved a categorical accuracy of 83.15%, with top-2 and top-3 accuracies of 91.36% and 95.34%, respectively, while weighted average precision, recall, and F1-score were reported as 0.89, 0.83, and 0.83. These results demonstrate the model's ability to classify a diverse range of lesion types effectively, though challenges remained with rarer lesion classes due to limited data. The authors highlighted the suitability of MobileNet for resource-constrained environments, such as mobile diagnostic systems, due to its efficient architecture and low computational demands [18].

These studies illustrate significant progress in applying deep learning to skin lesion classification. Models such as VGG16, MobileNet, and ResNet have shown diagnostic accuracy comparable to traditional methods. However, challenges remain in accurately identifying rare and complex lesions, improving model performance across diverse datasets, and reducing dependence on extensive and varied data collections.

This thesis aims to tackle these challenges by developing a customized model for practical applications. The objective includes enhancing the classification of rare and complex lesions, improving generalization across datasets, and overcoming limitations associated with small training datasets through advanced augmentation techniques. Special attention will be given to refining the classifications of slight characteristics in challenging cases, including small or irregular lesions.

The proposed system will be developed as an accessible online platform featuring a user-friendly interface, intended to assist both healthcare professionals and the general public. By incorporating advanced AI technologies into practical medical workflows, this project seeks to support early detection of skin cancer, optimize diagnostic processes, and contribute to improved patient outcomes.

2.3. Convolutional Neural Network (CNN) in Machine Learning.

Deep learning methods have become a critical component in the development of modern technology driven by the increasing need to analyse and extract meaningful patterns from visual datasets. Convolutional Neural Networks (CNNs) are widely recognized for their ability to handle tasks such as image classification and examining visual data. These networks employ structured layers designed to isolate critical features, effectively optimizing the model by reducing trainable parameters while maintaining high accuracy. Built with a series of layers, each uniquely designed to perform a specific operation, from detecting basic shapes to more complex classification tasks, CNNs ensure that important information is retained and refined throughout the training process. As a result, CNNs serve as a powerful foundation for analysing visual data, with applications ranging from object detection to medical imaging classification.

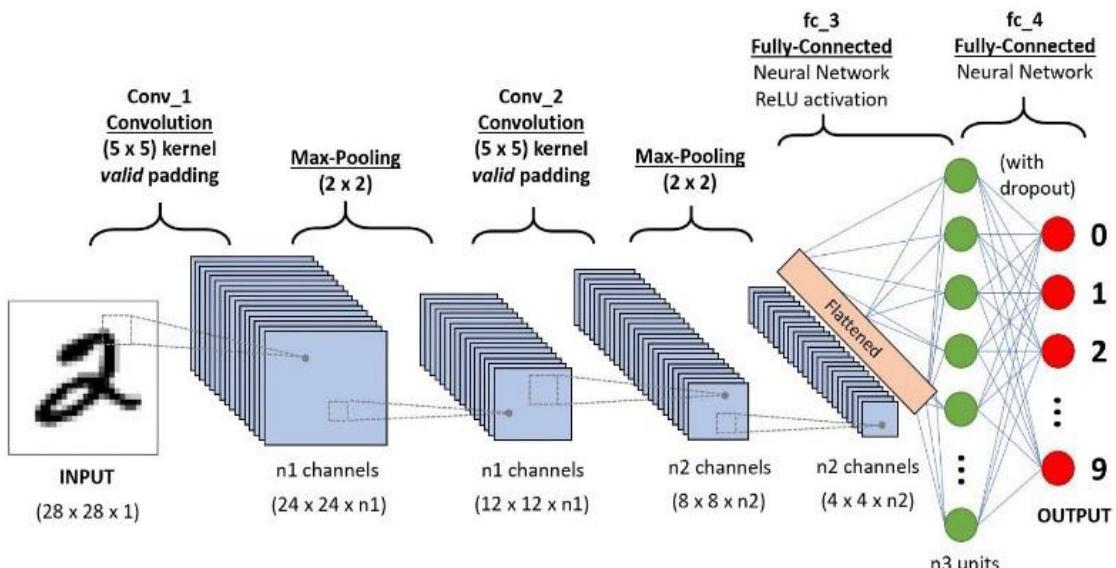


Figure 2.17: The overview of CNN architecture [29].

2.3.1. Overview of CNN architecture.

2.3.1.1. Input Layer.

The foundational layer serves as the starting point for processing raw information, which is commonly presented as 2D or 3D images. Its primary is to accept unprocessed input and forward it to later layers for deeper examination. These data are often structured as matrices and can represent either RGB (color) images or grayscale (black-and-white) images. This layer ensures that the model has access to the original visual data, laying the foundation for further feature extraction and analysis.

2.3.1.2. Convolutional Layer

The component responsible for convolution is crucial in enhancing the model's performance by identifying and processing key patterns in data. This is achieved by employing optimized matrix, often referred to as a filter, which moves systematically across the provided data. With every movement, the kernel calculates values for specific sections of the image, producing a feature map that highlights critical details of the input. During the training process, these filters are adjusted to identify patterns like edges, corners, or textures, enabling CNNs to analyse and handle complex visual information effectively [19], [20].

$$Z(i, j, k) = \sum_{m=0}^{F-1} \sum_{n=0}^{F-1} X(i + m, j + n) \cdot W(m, n, k) + b_k$$

Equation 2.1: The formula for convolutional layers [21].

Where:

$Z(i, j, k)$: Denotes the calculated result at coordinates k

$X(i + m, j + n)$: Denotes to the pixel intensity at the specified location in the input image matrix.

$W(m, n, k)$: Denotes the weight of the filter (or kernel) applied during the process

b_k : Bias associated with the kernel k , which helps adjust the output to better fit with the data

Once features are extracted through the convolutional layer, the next stage involves reducing data complexity and simplifying the resulting feature map. This operation reduces its spatial size, making the model less computationally expensive and less likely to overfitting. Pooling works by summarizing data from small sections of the feature map, enabling the model to concentrate on essential patterns while ignoring irrelevant details [19].

2.3.1.3. Pooling Layer.

After extracting features in the initial processing stage, the next step simplifies the feature representation. This process modifies the structure of the feature matrix, decreasing computational demands and minimizing the potential for overfitting [21]. Pooling operates by summarizing information from smaller regions within the feature map, enabling the model to concentrate on essential patterns while ignoring less relevant details. Two common pooling methods are widely used:

- **Maximum Pooling:** This method chooses the maximum value from a defined region within the feature map, ensuring that the most significant features are preserved. [20].

$$Z(i, j) = \max(X(i \cdot S : i \cdot S + F - 1, j \cdot S : j \cdot S + F - 1))$$

Equation 2.2: The formula for maximum pooling [22].

- **Average Pooling:** This method computes the average value across a region, which helps to smooth out features and reduce sensitivity to noise [20].

$$Z(i, j) = \frac{1}{F^2} \sum_{m=0}^{F-1} \sum_{n=0}^{F-1} X(i \cdot S + m, j \cdot S + n)$$

Equation 2.3: The formula for average pooling [23].

Where: F : The size.

S : The stride.

Pooling significantly impacts deep learning by improving both efficiency and effectiveness. Max pooling select the maximum activation within a specified area, allowing the model to focus on the most relevant attributes in the input data. Conversely, average pooling mitigates the impact of small variations by calculating the mean value in a defined area, enhancing the model's ability to adapt to new data. Additionally, pooling simplifies the feature structure, reducing the complexity of the model and lowering computational requirements. This technique optimizes the process while preserving essential information, leading to improved model stability and overall performance [21], [19].

2.3.1.4. Normalization Layer.

Normalization layers are vital components in deep learning architecture, where they help stabilize training and improve overall network performance. Among these, batch normalization is one of the most widely used techniques. It works by normalizing the outputs

of each layer using the mean and variance calculated from a mini-batch of data [25]. This process ensures consistent input distribution for subsequent layers, addressing challenges like vanishing and exploding gradients [24]. Additionally, trainable parameters such as gamma (γ) and beta (β) are applied post-normalization, allowing the model to adapt to different datasets and retain flexibility in learning complex patterns [25].

Batch normalization improves the training process by ensuring consistent gradient flow across layers, thereby accelerating convergence. This approach is a key innovation in modern neural network architectures, empowering them to effectively handle advanced applications, including image recognition and medical diagnostics, while enhancing both performance and reliability [24], [25].

$$\begin{aligned}\mu_B &= \frac{1}{m} \sum_{i=1}^m x_i \\ \sigma_B^2 &= \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \\ \hat{x}_i &= \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \\ y_i &= \gamma \hat{x}_i + \beta\end{aligned}$$

Equation 2.4: The formula batch normalization [25].

Where

μ_B : Represents the average value calculated for the mini-batch.

σ_B^2 : Refers to the variance determined for the mini-batch.

ϵ : A very small value to avoid division by zero.

2.3.1.5. Flattening Layer.

The flattening layer is a vital component in the prediction models, closing the gap between feature extraction and decision-making. Its primary function is to convert multi-dimensional data, such as 2D feature maps or 3D tensors, into a one-dimensional vector. This transformation ensures that the data is compatible with fully connected layers where each element of the vector represents a particular neuron in the subsequent layer. By retaining the key features extracted from previous layers, the flattening layer organizes the data in a format suitable for tasks like prediction and classification [26], [27].

In deep learning, particularly in image processing, convolutional and pooling layers often produce 2D or 3D output maps. The flattening layer compresses these complex representations into a single linear vector, simplifying the data structure and enabling the

network to effectively utilize the extracted features. This transition plays as an important role in linking the learned features to the decision-making process in fully connected layers [28].

2.3.1.6. Fully Connected Layer.

The last processing layer is an essential element in the structure of neural networks, serving as the final stage where all extracted features are integrated to make predictions. After data is processed through convolutional and pooling layers, the flattening layer converts it into a one-dimensional vector. The fully connected layer then establishes a connection between every input feature and each neuron in the output layer, creating a dense network similar to traditional neural networks. This comprehensive connectivity ensures that all the features learned during training contribute effectively to the decision-making process.

The primary role of the layer is to process information from previous layers into organized and meaningful output. By combining high-level features, it enables the network to perform tasks such as image classification or predictions task with precision. Its universal connection structure makes it indispensable for applications that require accurate and reliable results [26], [28].

$$Z(i,j) = \sum_{i=0}^n W_{ij} \cdot X_i + b_j$$

Equation 2.5: The formula for a fully connected layer [22].

- Where:
- Z_j :** Represents the output value at node j derived from linear combination of inputs and the bias term.
 - W_{ij} :** The weight connecting input i to node j , determining the input's impact on the output.
 - X_i :** Denotes the input value from the previous layer, serving as the information passed to the current layer for further computation.
 - b_j :** The bias term for node j , enabling independent adjustment of the output to improve training accuracy.
 - n :** The total counts of data points from the preceding layer, corresponding to the number of neurons involved in the calculation of the active node.

2.3.1.7. Activation Function.

This performs a vital function by incorporating non-linear behaviour into the model, enabling it to represent complex dependencies between input and output variables. This

approach allows the model to reveal detailed connections and multi-level frameworks. By applying activation functions after each convolutional operation, CNNs can efficiently address non-linear challenges, which is fundamental for tasks such as feature extraction and image recognition [20], [29].

- **ReLU (Rectified Linear Unit)**

This activation is widely used for its straightforwardness and high performance in deep learning models. The function converts all negative values to zero, while preserving positive values. This simple yet powerful transformation not only provides the necessary non-linearity to capture complex data relationships but also addresses the problem of diminishing gradients, often encountered during the training of large neural networks. Due to its efficiency and reliability, ReLU has become a fundamental element in most modern deep learning architectures. [30].

$$f(x) = \max(0, x)$$

Equation 2.6: The formula for ReLU [30].

- **Softmax**

Softmax is commonly employed in the concluding step to solve classification problems. It converts raw network outputs into probabilities, ensuring that their total adds up to one. This makes it highly effective for multi-class classification, as it simplifies the evaluation of model predictions. For example, in image recognition, the category with the greatest confidence is identified as the final result, ensuring precise and reliable results. [26].

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}}$$

Equation 2.7: The formula for Softmax [21].

2.3.2. DenseNet architecture.

Dense Convolutional Network, introduced in 2017, is a ground breaking architecture that revolutionized the way neural networks handle information flow and gradient propagation. Unlike traditional models, where connections are restricted to surrounding layers, DenseNet implements a dense connectivity mechanism. This means that each layer can directly access the outputs of all previous layers within the same block, ensuring better reuse of features and more efficient gradient flow. This design enhances training stability and allows for greater efficiency, particularly in deep networks where traditional architectures often struggle with vanishing gradients and inefficient feature utilization [31].

One of the key strengths of DenseNet is its adaptability. The architecture has been refined into multiple variants that balance computational cost and feature extraction abilities. Researchers can fine-tune parameters such as network depth and the degree of feature reuse to optimize DenseNet for specific tasks. DenseNet's versatility has been demonstrated in diverse applications, from image recognition to medical imaging, and even complex tasks such as image segmentation, where feature reuse and efficiency are critical [32]. By achieving an effective optimal balance between accuracy and computational effectiveness, DenseNet has become a popular solution for modern machine learning challenges [31].

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112 × 112		7 × 7 conv, stride 2		
Pooling	56 × 56		3 × 3 max pool, stride 2		
Dense Block (1)	56 × 56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56 × 56		1 × 1 conv		
	28 × 28		2 × 2 average pool, stride 2		
Dense Block (2)	28 × 28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28 × 28		1 × 1 conv		
	14 × 14		2 × 2 average pool, stride 2		
Dense Block (3)	14 × 14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14 × 14		1 × 1 conv		
	7 × 7		2 × 2 average pool, stride 2		
Dense Block (4)	7 × 7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1 × 1		7 × 7 global average pool		
			1000D fully-connected, softmax		

Figure 2.18: The comparison table of all DenseNet architectures [31].

2.3.2.1. Key features of the DenseNet model

2.3.2.1.1. Denseblock.

Dense Blocks are a fundamental part of the DenseNet architecture, designed to enhance feature learning and ensure efficient information flow in deep neural networks. Each layer in a Dense Block directly connects to all previous layers within the same block. This structure allows features learned in earlier layers to be readily available to later layers, minimizing redundancy and avoiding repeated computations. By utilizing this dense connectivity, the network effectively reuses learned features, leading to improved efficiency in training and faster convergence [31].

$$\mathbf{x}_l = \mathbf{H}_l([\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{l-1}])$$

Equation 2.8: The formula for Dense Block [31].

Where $[x_0, x_1, \dots, x_{l-1}]$: concatenation.

Dense Blocks offer a significant advantage by ensuring stable gradient flow during backpropagation. Through direct layer connections, they effectively resolve the problem of diminishing gradients, a frequent challenge in deep learning architectures [27]. This structure allows DenseNet to achieve superior performance with fewer parameters, making it a highly efficient computational solution [28]. Furthermore, Dense Blocks enhance feature reuse and optimized parameter efficiency, highlighting their importance in DenseNet's success across applications such as image recognition.

2.3.2.1.2. Transition Layer.

Transition layers play a critical role in DenseNet architectures by acting as connectors between Dense Blocks and managing the dimensions of feature maps and the count of channels. Without these layers, the network could become excessively large, making it computationally inefficient and difficult to optimize. Transition layers typically consist of two primary operations that enhance the network's efficiency and maintainability.

- The first operation involves the implementation of a convolutional layer with a 1×1 filter size. This operation reduces the count of feature channels in the feature maps, a technique known as feature compression. By focusing on preserving the most essential information while removing unnecessary details, this step minimizes redundancy and decreases the number of parameters, which ultimately speeds up the computation. This approach allows the network to maintain minimal size and ensures that only the most relevant data is passed to subsequent layers [24].
- The second step applies an average pooling layer with a 2×2 kernel. This reduces the structure of the feature maps, helping to lower computational demands without adding unnecessary complexity. In contrast to max pooling, which captures only the most dominant features, average pooling calculates the average of the values in each section, offering a more comprehensive view of the input data. This method is particularly useful for retaining important details that might otherwise be missed [33].

$$y = AvgPool(H_t(x))$$

Equation 2.9: The formula for Transition layer [27].

Where $H_t(x)$: The convolutional layer with a kernel size of 1×1 to reduce the dimensions of the output channels.

2.3.2.1.3. Growth Rate.

The rate is a crucial factor in DenseNet, determining the speed at which the number of features increases in each layer within a Dense Block. For DenseNet models designed for ImageNet, the growth rate is usually set to 32. A higher Growth Rate increases the amount of information added at each layer, enabling the network to learn richer and more detailed feature representations. However, this comes with the disadvantage of higher computational requirements and increased model complexity. On the other hand, a smaller growth rate decreases computational requirements and results in a more compact model, but it might restrict the network's potential to learn complex features [31]

$$\text{Number of features} = k_0 + l \cdot k$$

Equation 2.10: The formula for Growth Rate [31].

Where

k_0 : The starting set of output channels in the Dense Block.

l : The total depth of layers in the Dense Block.

2.3.3. Inception architecture.

The Inception module represents one of the most innovative advancements in deep learning, initially developed to optimize feature extraction in convolutional neural networks. This architecture introduced a new approach by incorporating parallel branches, which makes it possible for the network to manage information at different scales at the same time. This design not only improves computational efficiency but also improves the network's potential to understand sophisticated patterns, marking a significant evolution in deep learning techniques [35].

In addition to boosting performance, the Inception module is widely recognized for its ability to balance computational cost and learning capacity, making it suitable for addressing complex deep learning problems. Its efficiency and versatility have made it a cornerstone for advanced architectures, such as Inception-v4, Inception-v3 and Inception-ResNet, and have inspired further research into more efficient and scalable network designs. By connecting theoretical advancements with practical applications, the Inception module has inspired further development in designing more efficient and scalable networks, maintaining its influence on modern deep learning practices [36].

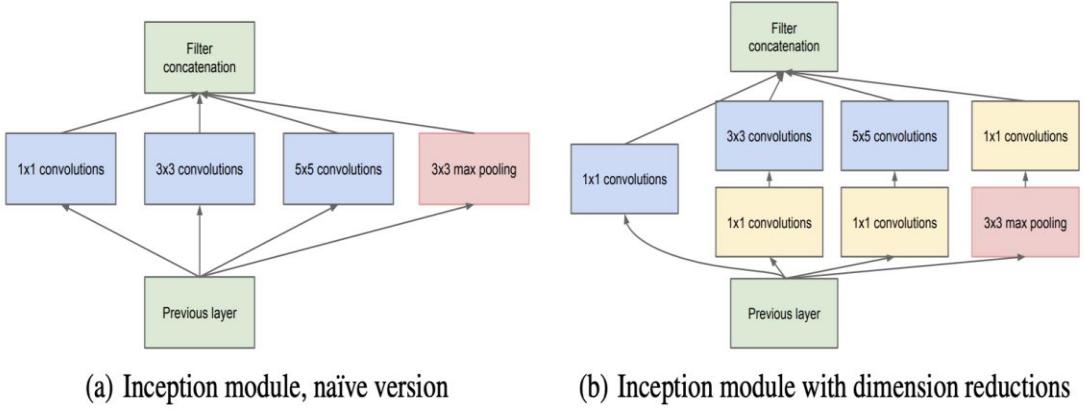


Figure 2.19: The Inception overall architecture [35].

2.3.3.1. Key features of the Inception model.

2.3.3.1.1. Multi-Scale Feature Extraction.

The Inception Module's ability to simultaneously extract information at many scales is one of its key advances. In order to accomplish this, parallel overlaid product filters of various sizes such as: 1x1, 3x3, and 5x5 are used in the same module [37]. Various degrees of detail are highlighted by each filter:

- **The 1x1 filters:** By reducing the dimensionality of the data and extracting pixel-level correlations, the 1x1 filter improves the representation of detail [34].
- **The 3x3 and 5x5 filters:** These capture features at a wide range of levels and extract patterns at medium and large scales [49]. Pattern recognition is made more flexible by the network's ability to preserve spatial systems without being constrained by a particular filter size thanks to its parallel construction [37].

Pattern recognition is made more flexible by the network's ability to preserve dimensional systems without being constrained by a particular filter size thanks to its parallel construction.

2.3.3.1.2. Dimension Reduction Through 1x1 Filter.

The Inception Module employs 1×1 convolutional filters as a dimensionality reduction technique to enhance computational efficiency. These filters are applied before the larger 3×3 and 5×5 filters to reduce the number of input data channels while retaining essential feature information. By employing this particular strategy, the module minimizes the computational

cost and parameter count, enabling efficient feature extraction and reducing the risk of memory saturation during training on large datasets [36].

2.3.3.1.3. Parallelism And Parallelism.

A key innovation of the Inception module lies in its parallel branch structure, which processes input data simultaneously across multiple paths. This design allows the network to specialize in learning different levels of abstraction and activates only the branches most relevant to the input data, promoting efficient computation and feature learning. By combining outputs from these parallel branches, the module creates a more comprehensive representation of the input data without redundancy, enhancing the overall learning ability of the network [38], [35].

2.3.3.1.4. Integrated Max-Pooling Layer

The Inception module also integrates max-pooling layers within its parallel branches to complement the convolutional filters. The pooling layer reduces the overall structure of the generated results, enhancing computational efficiency while retaining crucial feature representations. This process helps maintain stability under slight changes in the data, making the module powerful to minor noise or modifications in images [38]. Together, these innovations make the Inception module a versatile and efficient architecture for deep learning tasks.

2.4. Techniques for website implementation.

This project is developed on a web platform, with HTML (Hypertext Markup Language) acting as the foundation for defining the structure and content of online pages. HTML enables for the creation of essential elements such as text, photos, tables, links, and a variety of other elements that contribute to the general layout of the website [39]. CSS (Cascading Style Sheets) is used to format and alter the presentation of HTML elements, ranging from color, typeface, and size control to sophisticated and flexible layout arrangement [40].

Furthermore, React.js, a JavaScript package, is used extensively in the development of user interfaces (UI) for modern web applications. React.js offers an efficient and systematic approach to front-end application development, allowing developers to design interactive, dynamic, and maintainable interfaces. React.js is regarded as a leading solution for web application development, offering component reusability, superior performance through Virtual

DOM technology, and extensive flexibility to address the increasing complexity and diverse needs of modern users [41].

2.5. Techniques for machine learning models and API service.

This thesis combines a variety of advanced Python modules and tools for data processing, analytics, visualization, machine learning, and web construction, with each playing an important part in reaching study objectives.

Firstly, NumPy (Numerical Python), a basic library in scientific computing, offers high-performance multidimensional array objects as well as a comprehensive set of arithmetic algorithms. NumPy's widespread use in scientific, engineering, and data analysis communities illustrates its significance in computational research [42].

Secondly, Pandas high-performance data structures and versatility make it indispensable for activities like data cleaning process, manipulation, finding, and analysis. Pandas' tremendous features support the effective arrangement and preparation of data for future processing [43].

Thirdly, Seaborn and Matplotlib are used to visualize data in this thesis. Seaborn improves the beauty and understandability of statistical graphs by extending Matplotlib's fundamental capability. Meanwhile, Matplotlib provides excellent control over plots, enabling the creation of detailed and customized visualizations. [44],[45].

Last but not least, TensorFlow is a powerful, scalable, and adaptable software framework for development and training model. TensorFlow's rich ecosystem and outstanding performance make it the ideal tool for solving complicated mathematical issues [46].

To make web API development and deployment easier, Ngrok is utilized, which provides options for constructing secure channels and providing public access to local web services. This tool is particularly useful for evaluating and implementing API features in research situations [47].

CHAPTER 3

METHODOLOGY

3.1. Proposed method for skin lesions prediction.

Deep learning has been used to create an effective skin lesion classification framework capable of accurately distinguishing between cancerous and non-cancerous lesions. This approach is designed to be efficient and accurate in classification.

This deep learning-based framework efficiently classifies skin lesions into cancerous and non-cancerous categories with high accuracy. The process begins with image preprocessing, including resizing (224×224), label encoding, and normalization. The dataset is then divided into training, validation, and test subsets. To improve model generalization and handle class imbalance data augmentation is applied.

Four state-of-the-art convolutional neural network architectures—DenseNet201, InceptionV3, Xception, and ViT—are employed for feature extraction. Model training is optimized through hyperparameter tuning, enhancing classification accuracy. After training, model evaluation assesses its performance before making final predictions.

This framework is intended to detect seven distinct forms of skin lesions, both malignant and non-cancerous. The combination of data pre-processing, deep learning algorithms, and model optimization has resulted in a sophisticated system that can help with enhancing outcomes through early detection of skin diseases.

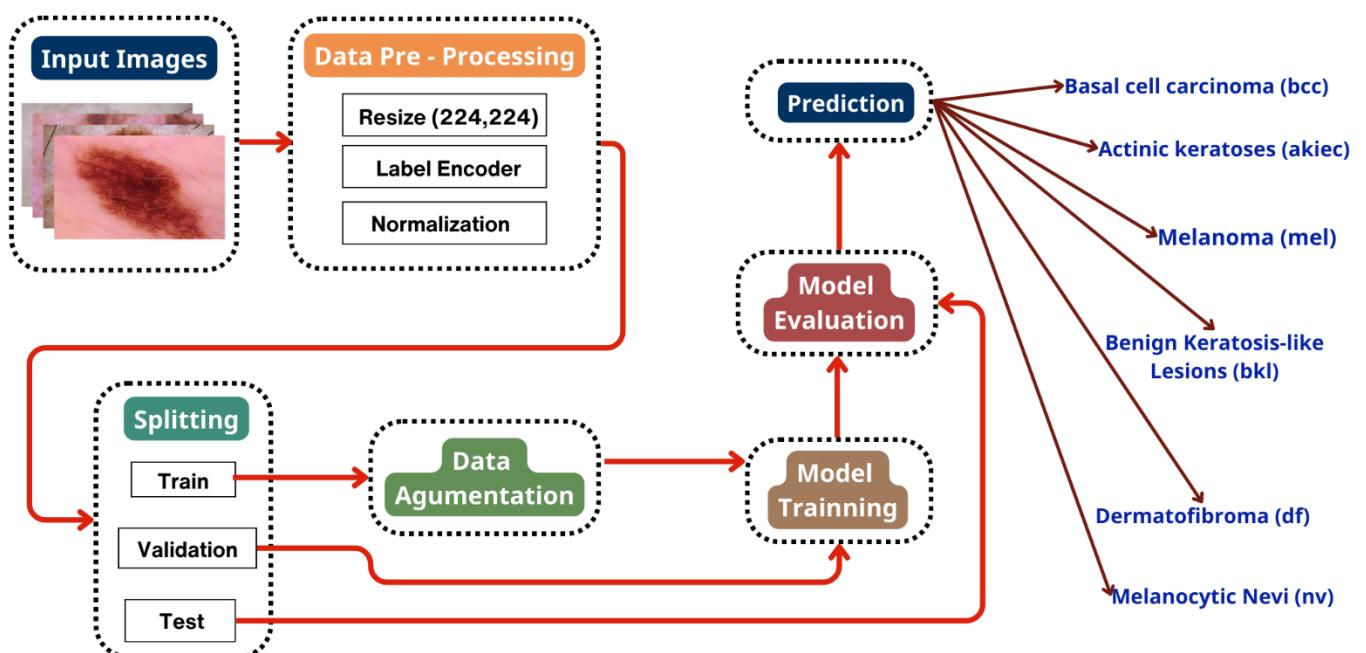


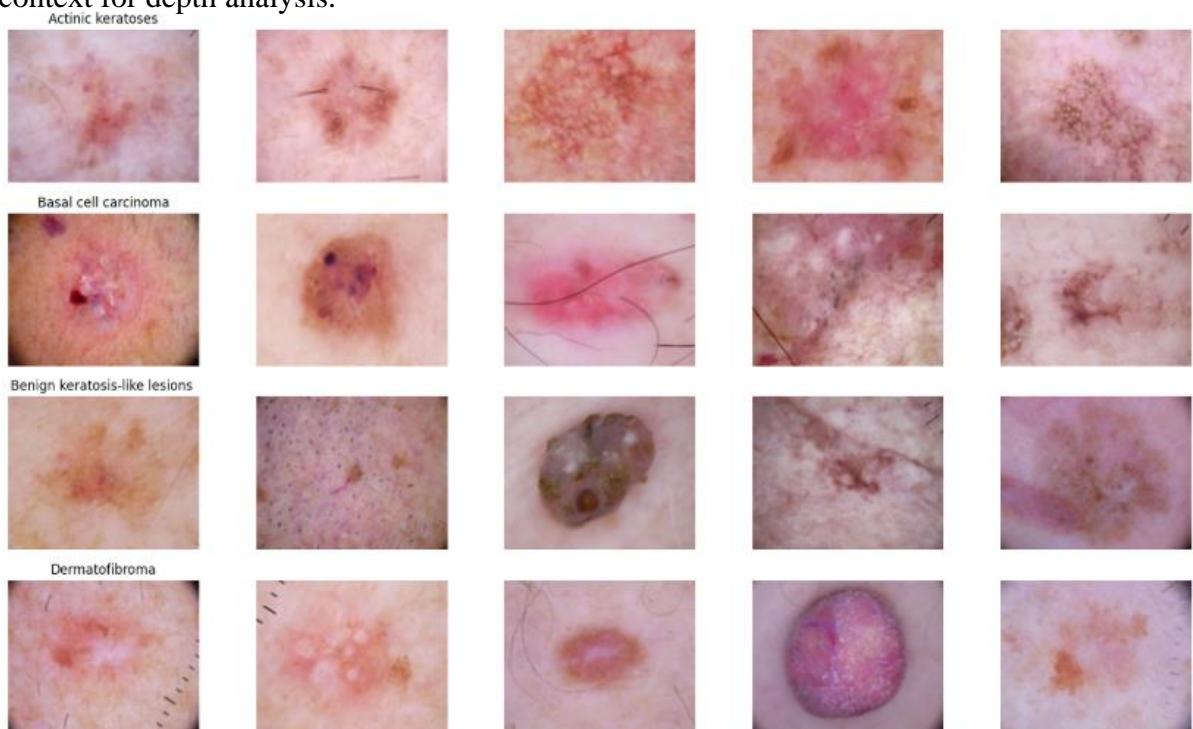
Figure 3.1: Proposed methodology in this study.

3.2. Dataset.

The HAM10000 medical dataset includes a collection of 10,015 skin lesion images, organized into seven unique classification categories, including both benign and malignant cases. This dataset stands out as one of the most comprehensive and varied resources available for dermatological diagnosis using deep learning. The images were collected from multiple sources, predominantly dermatology clinics and specialized dermoscopy applications, ensuring a wide diversity of skin lesion forms in everyday medical practice, from common to rare cases [48].

Each image within the dataset has been carefully annotated by medical experts, representing seven lesion categories. Among these, melanocytic nevi (nv) is the most frequently occurring category, typically marking as benign pigmented lesions. In contrast, the melanoma (mel) category identifies potentially life-threatening malignant lesions, highlighting the importance of early detection to prevent cancer development. Other categories include benign keratosis-like lesions (bkl), basal cell carcinoma (bcc), and actinic keratoses or intraepithelial carcinoma (akiec), which range from precancerous to cancerous states. Additionally, less prevalent types such as dermatofibroma (df) and vascular lesions (vasc) are included to ensure comprehensive representation [48].

The dataset offers not only a diverse range of lesion types but also high-quality images with variations in resolution, lighting, and skin tone, enabling deep learning models to capture essential features for accurate classification. Furthermore, each image is accompanied by metadata, such as patient age, gender, and lesion location on the body, which provides valuable context for depth analysis.



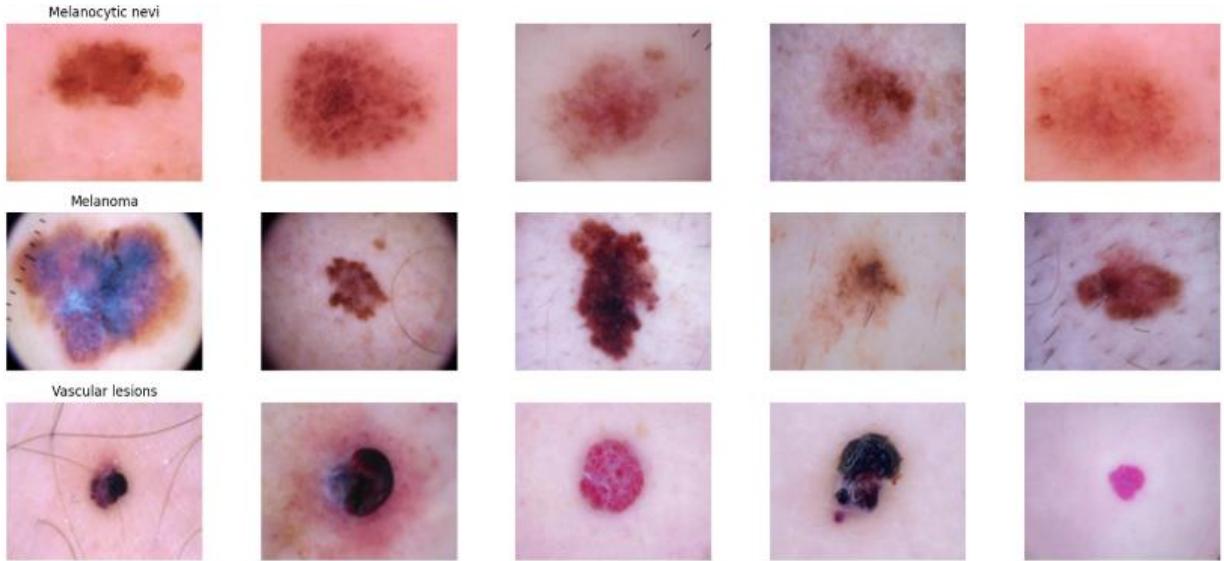


Figure 3.2: The examples of sample lesions in the dataset [55].

3.3. Data visualisation.

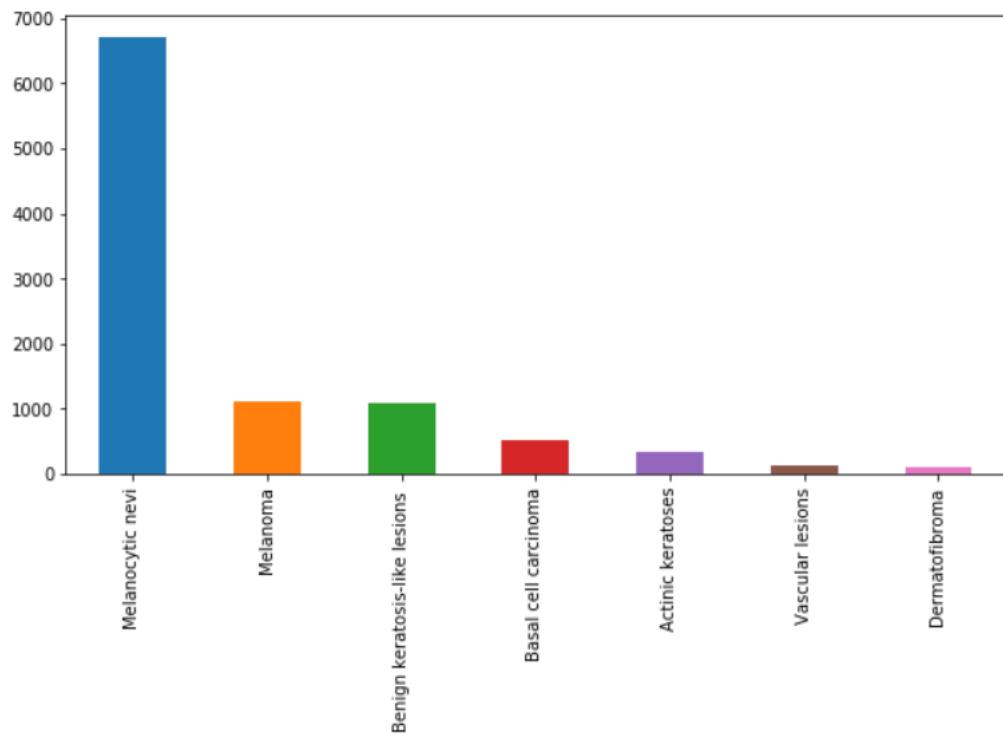


Figure 3.3: The bar chart illustrates how skin lesion types are distributed within the dataset.

The chart shows the distribution of seven types of skin lesions, with melanocytic nevi being the most common, making up about 7000 cases, far more than the others. Melanoma and benign keratosis-like lesions have similar counts, ranging from 1000 to 1500 cases. In comparison, basal cell carcinoma, actinic keratoses, vascular lesions, and dermatofibroma are much less common, each with fewer than 1000 cases. This imbalance in the dataset reflects the

fact that some skin lesions are more frequent in the general population, which creates challenges for developing accurate diagnostic models.

The unequal distribution of cases causes machine learning models to focus more on common lesion types while performing poorly on rare ones. This is particularly concerning in healthcare, where rare but serious conditions like melanoma might be missed, leading to severe consequences. To build reliable diagnostic models, balanced data is essential, as imbalances reduce the model's ability to accurately detect uncommon types of lesions, affecting its performance and real-world application.

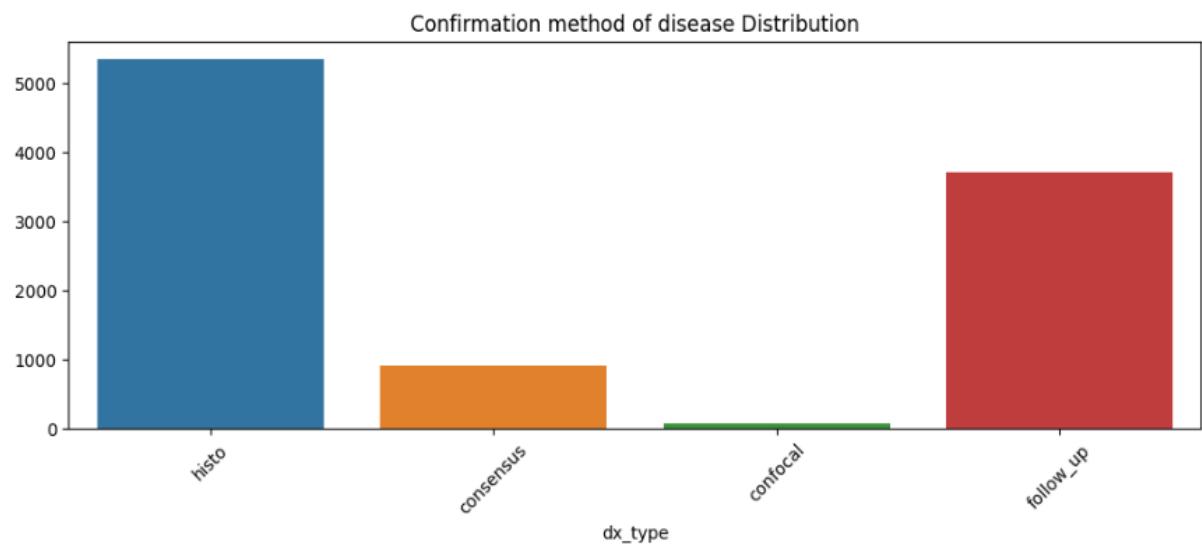


Figure3.4: The bar chart shows the distribution of diagnostic confirmation methods for diseases.

The chart highlights that histopathology (Histo) is the most commonly used method for diagnosing diseases, with over 5000 cases reported. This method is widely valued for its accuracy, as it involves a detailed examination of tissue samples. Its consistent use emphasizes its reliability and its critical role in accurately identifying malignant lesions, making it a preferred choice in medical diagnostics.

The second category is Follow-up, which has almost 4000 instances, demonstrating the necessity of clinical follow-up in cases where biopsy is not required. This is evidenced by the frequent use of follow-up for benign lesions, such as nevus, due to its ease and minimal invasiveness. Meanwhile, consensus and confocal have very small numbers, with consensus accounting for less than 1000 occurrences.

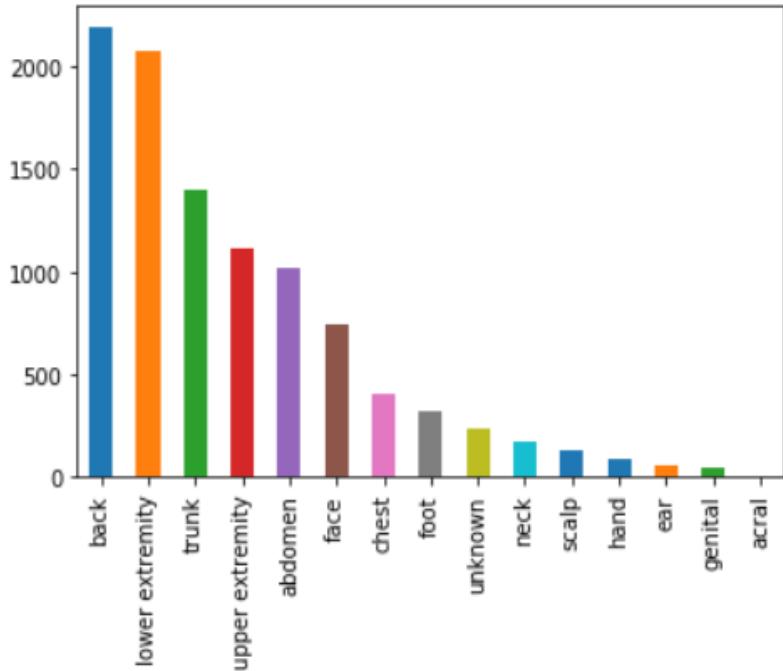


Figure 3.5: The bar chart displays the counts of the body areas that have skin lesions.

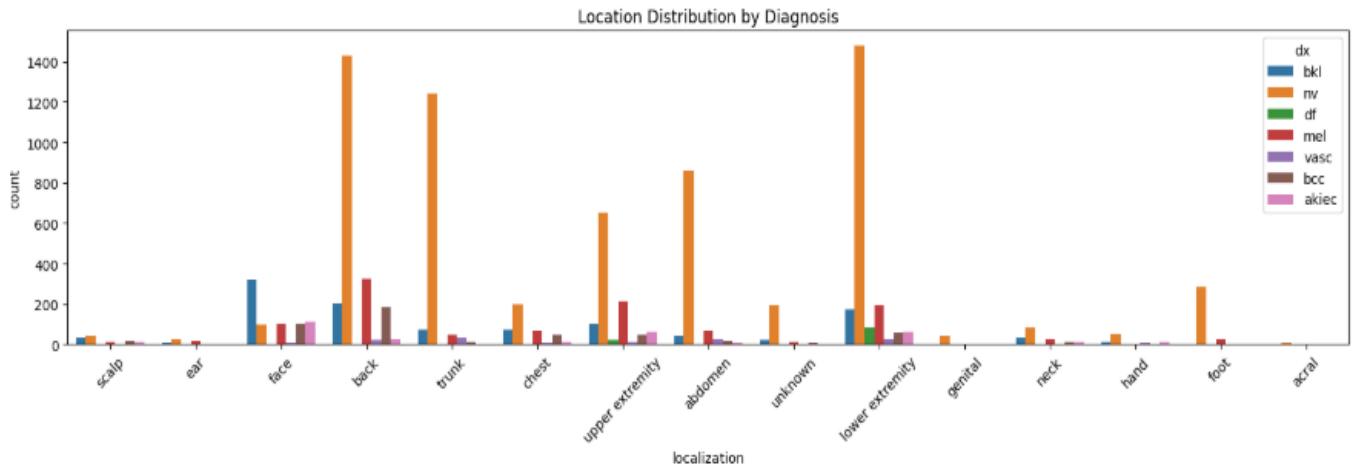


Figure 3.6: The bar chart presents the distribution of all skin lesions types across each body area.

The two charts above describe the distribution of skin lesions by body location, as well as the types of lesions associated with each. The first graph shows the overall number of skin lesions in different places. The results revealed that the back had the most skin lesions, with over 2000 cases, followed by the lower extremity and trunk. These are all places with big skin areas that are frequently exposed to sunlight, which is one of the leading causes of skin damage. The abdomen, cheeks, and chest also had a high number of lesions. Meanwhile, smaller or less apparent skin areas, such as the ears, neck, genitals, and soles, have a high density of lesions.

The second chart explores deeper into the danger of skin injury types across each body area. The findings revealed that common lesions, such as melanocytic nevi (nv) and benign keratosis-like lesions (bkl), appeared primarily on the back, trunk, and lower extremities, with

reduced presence in other places like the face and belly. Melanoma (mel) - a deadly type of skin cancer - is most commonly observed on the back and lower limbs, indicating the extent of the damage in these locations. Vascular lesions (vasc) are uncommon and primarily occur on the face and hands. The ears, throat, and genitals have fewer forms of lesions, as seen by the very low number in these areas.

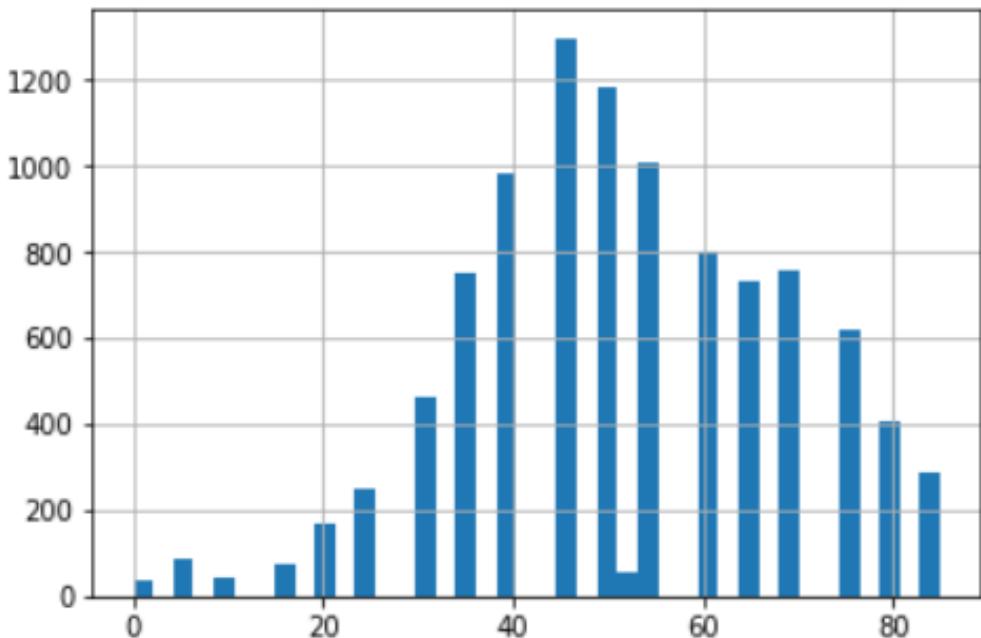


Figure 3.7: The frequency chart above presents the age distribution of all patients.

The frequency chart above presents the age distribution of skin lesions. According to data, adults aged 40 to 60 have the highest number of instances, with a peak of roughly 1,200 in middle age. The younger (under 20) and older (over 80) age groups have much smaller numbers. This distribution represents the frequency with which skin damage occurs, particularly serious kinds such as skin cancer which happened more frequently in middle-aged and older adults.

3.4. Data pre-processing and Data augmentation.

Preparing data is a core part of the process in creating models designed for medical image analysis and diagnosis. It helps assure the model's performance and accuracy. This technique has two primary components: data pre-processing and data augmentation. In this work, data cleaning and data enhancement processes were used in depth to improve data quality and overcome data imbalance.

3.4.1. Data Pre-processing.

Cleaning data serves as an initial step in preparing and organizing images before they are input into the model. Firstly, all images are standardized to be the same size and format. This ensures an evenly distributed pixel structure, which helps to reduce errors during training. The image labels are then thoroughly verified to ensure that there are no problems, such as missing or inaccurate labels. Any inaccuracies in the labels are corrected or removed to ensure the dataset accurately reflects the characteristics of the lesion classifications.

3.4.2. Data Augmentation techniques.

To tackle the problem of data imbalance among skin lesion categories, approaches for generating more samples are used to boost the dataset size in underrepresented classes. Many other data augmentation techniques have been implemented using the imgaug library, such as:

- **Horizontal Flip:** The code flips the image from left to right. In a horizontal flip, the image is mirrored along its vertical axis, so what was originally on the left side of the image appears on the right side, and vice versa. This is useful because lesions might appear in different positions when images are captured. By including horizontal flips, the model learns that a lesion can look the same even if the image is mirrored. This technique helps the model focus on the features of the lesion rather than on its position in the image.
- **Vertical Flip:** Vertical flipping means that the image is flipped from top to bottom. Although it is less common in real-world scenarios for medical images to be upside down, including vertical flips forces the model to recognize lesion features regardless of the image's vertical orientation. This augmentation helps ensure that the model does not become too dependent on a specific "upright" view of the lesion. As a result, it can handle cases where the image might be taken in an unusual orientation.
- **Rotation (up to $\pm 90^\circ$):** The code randomly rotates the image by up to 90 degrees in either direction. This means that each image might be rotated anywhere between -90 degrees (counter clockwise) and +90 degrees (clockwise). Rotation is very important because images are not always captured perfectly aligned. By training on rotated images, the model learns that the lesion remains the same even if it is viewed from different angles. This rotation helps the model become less sensitive to the specific angle at which a lesion is photographed.

- **Zoom ($\pm 10\%$):** Zooming changes the size of the image by scaling it up or down. In this case, the image is zoomed in or out by 10%. Zooming in simulates a scenario where the lesion appears larger because the camera is closer, while zooming out simulates the lesion being smaller due to a greater distance from the camera. This augmentation helps the model learn to recognize lesions that appear in various sizes. It also helps the model adapt to images where the lesion occupies different proportions of the frame.
- **Width and Height Shift ($\pm 10\%$):** Shifting the image means moving it horizontally or vertically by a small amount. The code shifts the image by up to 10% of its width (for horizontal shifts) or its height (for vertical shifts). This type of transformation is useful because, in many real-life images, the lesion may not be perfectly centered. By training on images that are shifted slightly to the side or up and down, the model learns to recognize lesions even when they are not in the exact center of the image. This helps in making the model robust to slight misalignments that might occur during image capture.

After completion the data augmentation process focuses on low-volume lesion classes. Several photos from these classes are randomly selected and enhanced to produce new versions. Following improvement, the image is saved on the dataset and properly labelled based on its class, ensuring straightforward classification and retrieval

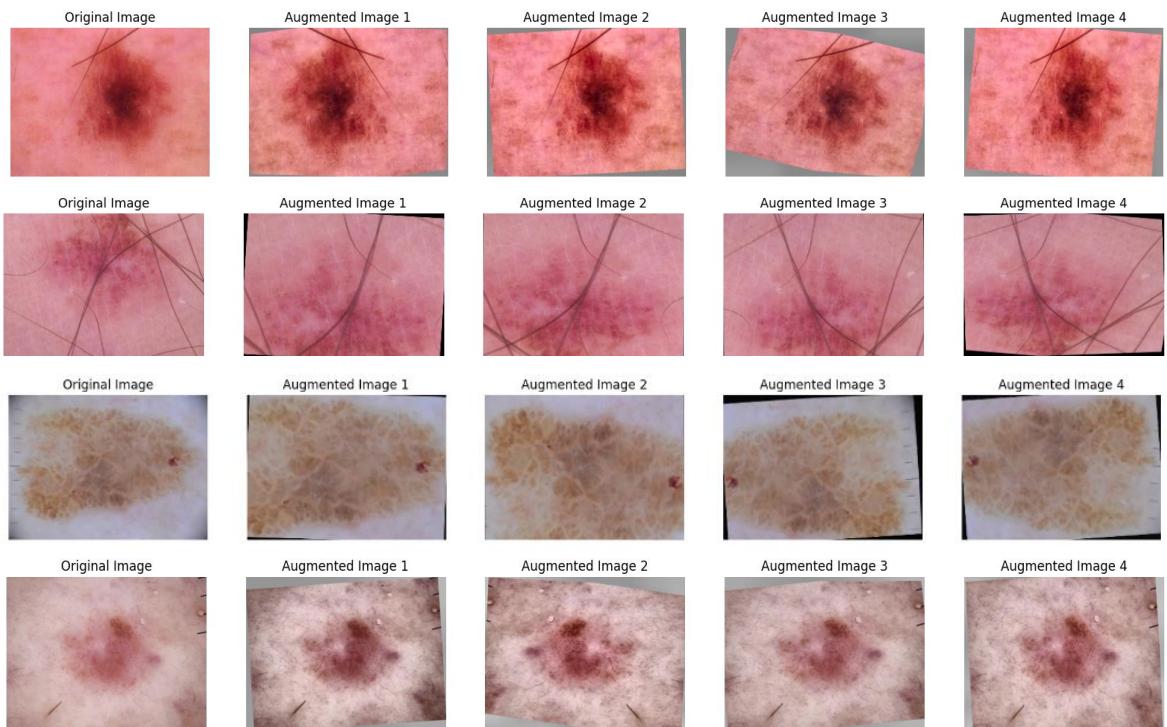


Figure 3.8: Examples of data augmentation techniques on the skin lesions dataset.

```

# Separate features (X) and target variable (y)
X = meta_data.drop(columns=['Updated_Diagnosis_Label'])
y = meta_data['Updated_Diagnosis_Label']

# Display the shape of X and y
print("Shape of features (X):", X.shape)
print("Shape of target variable (y):", y.shape)

→ Shape of features (X): (10015, 9)
Shape of target variable (y): (10015,)

[ ] from tensorflow.keras.preprocessing.image import ImageDataGenerator
# Data augmentation and generators
train_datagen = ImageDataGenerator(
    rescale=1./255,
    brightness_range=[0.8, 1.2],
    rotation_range=40,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True,
    fill_mode='nearest'
)
test_datagen = ImageDataGenerator(rescale=1./255)

train_generator = train_datagen.flow_from_dataframe(
    dataframe=X_train,
    x_col='Image_path',
    y_col='Updated_Diagnosis',
    target_size=(224, 224),
    batch_size=32,
    class_mode='categorical'
)
val_generator = test_datagen.flow_from_dataframe(
    dataframe=X_val,
    x_col='Image_path',
    y_col='Updated_Diagnosis',
    target_size=(224, 224),
    batch_size=32,
    class_mode='categorical',
    shuffle=False
)

test_generator = test_datagen.flow_from_dataframe(
    dataframe=X_test,
    x_col='Image_path',
    y_col='Updated_Diagnosis',
    target_size=(224, 224),
    batch_size=32,
    class_mode='categorical',
    shuffle=False
)

```

Figure 3.9: The code of data augmentation technique code.

In the ViT model, various data augmentation techniques were applied to the training dataset to enhance the model's ability to generalize effectively to new images. These techniques included random rotation (up to 90 degrees) to introduce different orientations of the same image, sharpness adjustment to simulate variations in image quality, and random horizontal flipping with a 50% probability to introduce further diversity in image representation. The validation and test datasets were not subjected to extensive augmentation but were processed using only resizing and normalization validation set to ensure consistent model evaluation. The dataset was stored in the Hugging Face Dataset format, making it easier to integrate with transformer-based deep learning models.

```
# Transformation for training data
train_tf = Compose([
    Resize((image_size, image_size)), # Resize the image to the specified size (e.g., 224x224)
    RandomRotation(90), # Randomly rotate the image by up to 90 degrees
    RandomAdjustSharpness(2), # Randomly adjust image sharpness with a sharpness factor of 2
    RandomHorizontalFlip(0.5), # Apply a horizontal flip with a 50% probability
    ToTensor(), # Convert the PIL image to a Tensor (pixel values range from 0 to 1)
    normalize # Normalize the image using `image_mean` and `image_std`
])

[] val_tf = Compose([
    Resize((image_size, image_size)), # Resize the image to the specified size (e.g., 224x224)
    ToTensor(), # Convert the PIL image to a Tensor (pixel values range from 0 to 1)
    normalize # Normalize the image using `image_mean` and `image_std`
])

[ ] def train_transforms(examples):
    examples['pixel_values'] = [train_tf(image.convert("RGB")) for image in examples['image']]
    return examples

def val_transforms(examples):
    examples['pixel_values'] = [val_tf(image.convert("RGB")) for image in examples['image']]
    return examples
```

Figure 3.10: The code for ViT augmentation.

3.5. Deep learning skin lesion predictions models.

Deep Learning has advanced rapidly over the past decade, offering innovative solutions to complex problems, especially in image processing and computer vision. Researchers have continuously improved model architectures to enhance feature extraction and address challenges like gradient vanishing, slow convergence, and high computational costs in training deep networks. This study employs four widely recognized models: DenseNet201, InceptionV3, Xception, and ViT. These models represent innovative advancements in deep learning, designed to maximize feature learning efficiency and computational performance, making them well-suited for image analysis tasks.

3.5.1. DenseNet201 model.

DenseNet201 was selected for this study due to its exceptional ability to optimize parameter usage and support the smooth flow of features, which are essential for analysing medical image datasets. Specially in medical imaging, the restricted access to labeled datasets remain a core challenge for advancing research and practice, making it crucial to use a model that can maximize learning efficiency. DenseNet201 achieves this by reusing features across layers, greatly lowering the number of parameters compared to standard convolutional neural networks. This approach minimizes memory requirements and at the same time it enhances the model's generalization ability, which is particularly beneficial when working with smaller datasets. Furthermore, its abilities to overcome the vanishing gradient challenge ensures stable training, even in deep architectures. These attributes make DenseNet201 a compelling choice for tasks that require capturing detailed features in images, aligning with the goals of this research [31].

3.5.2. InceptionV3 model.

InceptionV3 was chosen for its proven balance between computational efficiency and high performance, especially in tasks requiring analysis of high-resolution images. One of its key advantages lies in its modular architecture, which uses Inception modules to extract features at multiple scales within a single layer. This capability is particularly relevant in skin lesion classification, where lesions can vary significantly in size, shape, and texture. Additionally, InceptionV3 incorporates transformed convolutions to reduce the computational overhead, making it more feasible to train the model on standard hardware setups. The success it has shown in large image classification tasks makes it particularly well-suited for this study. By employing InceptionV3, the research aims to use its powerful feature extraction abilities to identify hidden patterns in medical images [38].

3.5.3. Xception model.

Xception was chosen for its innovative application of depth wise separable convolutions, which improve computational efficiency while maintaining accuracy. This design splits the convolution process into two stages: depthwise convolution, which processes each channel individually, and point-wise convolution, which combines the outputs from all channels. In addition, this design effectively reduces the parameter size and computational complexity in comparison to conventional convolutional approaches. Xception's ability to extract high-quality features has been demonstrated in tasks involving large and diverse datasets, making it a suitable choice for skin lesion classification, where precision is critical.

By incorporating Xception into this research, the study aims to take advantage of its architecture to achieve accurate and efficient analysis of medical images [49].

3.5.4. The Vision Transformer model.

The Vision Transformer is highly effective for skin cancer classification, as it analyzes images globally rather than relying on local features like traditional CNNs. By dividing an image into patches and processing them as a sequence, ViT captures critical variations in shape, texture, and color, essential for distinguishing between benign and malignant lesions. Its self-attention mechanism enhances accuracy by focusing on key regions, while its ability to handle high-resolution medical images preserves fine details crucial for diagnosis. Pretrained on large datasets, ViT performs well even with limited labeled medical data, and with data augmentation, it adapts to different skin types, lighting conditions, and lesion variations, reducing bias and improving classification accuracy. This adaptability, combined with its strong feature extraction abilities, makes ViT a powerful tool for automated skin cancer detection, often outperforming traditional deep-learning models. In this approach, a pretrained ViT model was chosen for its ability to detect subtle differences in medical images, with label-to-ID and ID-to-label mappings extracted to ensure accurate classification tasks.

3.5.5. Performance metrics.

Evaluating how well machine learning models perform is a vital process to confirm their suitability for solving specific tasks. This evaluation relies on performance metrics that offer detailed insights into how well the model operates, particularly in handling imbalanced data, generating reliable predictions, and generalizing effectively. For this study, several metrics were selected to evaluate the skin lesion classification models, each addressing distinct aspects of performance.

- **Accuracy** This metric provides a measure of how accurately the model predicts instances relative to the total outputs it produces [50]. While it provides an overall perspective, its effectiveness diminishes in scenarios with imbalanced datasets. For example, in a dataset where malignant cases are uncommon, a model that mostly predicts benign outcomes may show high accuracy but miss important malignant cases.
- **Sensitivity:** Also referred to as Recall, measures the model's effectiveness in identifying true positive cases [51]. This metric is especially crucial in medical applications, such as detecting malignant skin lesions, where missing even a single positive case could lead to severe consequences. By emphasizing sensitivity, the study ensures that the

model prioritizes capturing malignant cases, even if it increases the number of false positives.

- **Specificity:** This evaluation parameter emphasizes the model's ability to accurately examine true negative cases, thereby minimizing false positives [52]. In healthcare, this metric plays a key role in reducing unnecessary interventions, such as avoidable biopsies or surgeries. Balancing sensitivity and specificity help the model correctly identify serious cases while reducing the chance of overdiagnosis.
 - **Precision:** This evaluates the percentage of correctly identified positives among all predicted positive outcomes [51]. High precision is essential in clinical settings to ensure the trustworthiness of the model. A model with low precision could generate frequent false alarms, weakening confidence in its predictions and potentially leading to overdiagnosis.
 - **The F1-Score:** Combines both sensitivity and precision into a single metric, making it particularly useful for imbalanced datasets [51]. Moreover, the score ensures a detailed assessment of the model's capacity to detect malignant cases while ensuring the consistency and reliability of its predictions, which is essential for this research.
- .

CHAPTER 4

IMPLEMENT AND RESULTS

4.1. Experimental setup.

To address the multiclass classification problem, four advanced convolutional neural network models are used in the experiment: DenseNet201, InceptionV3, InceptionResNetV2, and Xception. These models, which are pre-trained, not only provide powerful feature extraction abilities, but also lay the groundwork for fine-tuning for a specific situation. The necessary modifications were made to guarantee that each model performed optimally in the testing environment.

After extracting features from the base model, a global average pooling layer is used to reduce the dimensionality of the data. Next, a fully connected layer containing 128 units is added, using the ReLU activation function to introduce non-linear behaviour. To address overfitting, a dropout layer with a 0.5 rate is applied before the final layer. The output layer includes 7 units with a softmax activation function, which represents the predicted probabilities for each category. Details of the layer structure are provided in Table 1.

Backbone	Type	Output Shape	Parameters
DenseNet201	(Base Model)	(None, 7, 7, 1920)	Pretrained
	GlobalAveragePooling2D	(None, 1920)	0
	Dense (128 units, ReLU)	(None, 128)	245,888
	Dropout (Rate: 0.5)	(None, 128)	0
	Dense (7 units, Softmax)	(None, 7)	903
InceptionV3	(Base Model)	(None, 5, 5, 2048)	Pretrained
	GlobalAveragePooling2D	(None, 2048)	0

	Dense (128 units, ReLU)	(None, 128)	262,272
	Dropout (Rate: 0.5)	(None, 128)	0
	Dense (7 units, Softmax)	(None, 7)	903
Xception	(Base Model)	(None, 7, 7, 2048)	Pretrained
	GlobalAveragePooling2D	(None, 2048)	0
	Dense (128 units, ReLU)	(None, 256)	262,272
	Dropout (Rate: 0.5)	(None, 256)	0
	Dense (7 units, Softmax)	(None, 7)	1.799

Table 4.1: The setup details for four neural network models trained in this research.

The models were optimized using the Adam optimization algorithm, with categorical cross-entropy as the loss function, a standard choice for multiclass classification. Accuracy was selected as the primary metric to assess the model's performance during training. To handle class imbalance, class weights were assigned based on the frequency of each category, ensuring fair representation across the dataset.

The training was performed over 10 epochs with a batch size of 100, incorporating normalized and augmented data to enhance the models' generalization abilities. The process was monitored using callbacks, including ReduceLROnPlateau, Early Stopping, and ModelCheckpoint. ReduceLROnPlateau adjusted the learning rate by 0.1 whenever the test loss showed no improvement for two consecutive epochs. Early Stopping stopped the training after three consecutive epochs with no progress and restored the weights of the best-performing model, ensuring optimal validation performance.

```

import tensorflow as tf
from tensorflow.keras.applications import DenseNet201, InceptionV3, InceptionResNetV2, Xception
from tensorflow.keras.layers import GlobalAveragePooling2D, Dense, Dropout
from tensorflow.keras.models import Model
from tensorflow.keras.optimizers import Adam
from tensorflow.keras.callbacks import ModelCheckpoint, ReduceLROnPlateau, EarlyStopping

# Define a function to build the model
def create_model(base_model, dense_units=128):
    """
    Build the model using a base model and custom fully connected layers on top.

    Args:
        base_model: A pre-trained base model without the top layers.
        dense_units (int): Number of units in the dense layer.

    Returns:
        model: A complete Keras model with added dense and dropout layers.
    """
    x = base_model.output
    x = GlobalAveragePooling2D()(x) # Add a global average pooling layer
    x = Dense(dense_units, activation='relu')(x) # Add a dense layer with specified units and ReLU activation
    x = Dropout(0.5)(x) # Add a dropout layer to reduce overfitting
    predictions = Dense(7, activation='softmax')(x) # Output layer with 7 units and softmax activation
    model = Model(inputs=base_model.input, outputs=predictions) # Define the final model
    return model

# List of base models
base_models = {
    "DenseNet201": DenseNet201(weights="imagenet", include_top=False, input_shape=(224, 224, 3)),
    "InceptionV3": InceptionV3(weights="imagenet", include_top=False, input_shape=(224, 224, 3)),
    "Xception": Xception(weights="imagenet", include_top=False, input_shape=(224, 224, 3))
}

# Loop through each base model
for model_name, base_model in base_models.items():
    print(f"Building and training model: {model_name}")

    # Freeze the layers in the base model
    for layer in base_model.layers:
        layer.trainable = False

    # Use 256 units for Xception, otherwise default to 128 units
    dense_units = 256 if model_name == "Xception" else 128

    # Build the complete model
    model = create_model(base_model, dense_units=dense_units)

    # Compile the model
    model.compile(
        loss='categorical_crossentropy',
        optimizer=Adam(),
        metrics=['accuracy']
    )

    # Define callbacks
    checkpoint = ModelCheckpoint(
        filepath=f'{model_name}_weights.h5', # Save model weights
        monitor='val_loss',
        save_best_only=True,
        verbose=1
    )
    reduce_lr = ReduceLROnPlateau(
        monitor='val_loss',
        factor=0.1,
        patience=2,
        min_lr=1e-5,
        verbose=1
    )
    early_stopping = EarlyStopping(
        monitor='val_loss',
        patience=3,
        restore_best_weights=True,
        verbose=1
    )

```

```

    # Train the model
    history = model.fit(
        train_generator, # Training data generator
        epochs=10, # Number of epochs
        batch_size=100, # Batch size
        validation_data=val_generator, # Validation data generator
        callbacks=[checkpoint, reduce_lr, early_stopping], # Callbacks
        class_weight=class_weights # Class weights
    )

    print(f"Finished training model: {model_name}")

```

Figure 4.1: The code for three deep learning models.

For the Vit model, a special training strategy was employed to optimize model performance. The key hyperparameters included a learning rate of 3e-5 using a cosine learning rate scheduler, a weight decay of 0.1 to prevent overfitting, a batch size of 64 per device, and 10 training epochs. The AdamW optimization algorithm was used to update model parameters efficiently. Model performance was assessed at the end of each epoch, and early stopping was enabled to ensure the best-performing checkpoint was selected based on validation accuracy. To streamline the training and evaluation process, Hugging Face's Trainer API was utilized, allowing for efficient data handling and logging of performance metrics.

```

args = TrainingArguments(
    f"{model_name}", # Path to save the model, dynamically set
    remove_unused_columns=False, # Whether to remove unused columns; default is False (keeps all columns)
    eval_strategy="epoch", # Evaluation strategy: evaluate at the end of each epoch
    save_strategy="epoch", # Save strategy: save the model at the end of each epoch
    learning_rate=3e-5, # Learning rate: step size for updating model parameters during training
    lr_scheduler_type="cosine", # Learning rate scheduler: "cosine" (cosine annealing), reduces LR gradually
    auto_find_batch_size=True, # Automatically find an appropriate batch size based on available memory
    per_device_train_batch_size=64, # Training batch size per device (e.g., per GPU)
    per_device_eval_batch_size=64, # Evaluation batch size per device
    weight_decay=0.1, # Weight decay for preventing overfitting (L2 regularization)
    num_train_epochs=10, # Total number of training epochs
    load_best_model_at_end=True, # Load the best model at the end of training
    metric_for_best_model="accuracy", # Metric to determine the best model (here, accuracy)
    logging_strategy="epoch", # Log loss at the end of each epoch
    report_to="none" # Disable reporting to external systems (e.g., WandB, TensorBoard)
)

```

```

[ ] trainer = Trainer(
    model,
    args,
    train_dataset=train_dataset,
    eval_dataset=val_dataset,
    tokenizer=image_processor,
    compute_metrics=compute_metrics,
    data_collator=collate_fn,
)

```

```

[ ] import torch
def collate_fn(examples):
    pixel_values = torch.stack([example["pixel_values"] for example in examples])
    labels = torch.tensor([example["label"] for example in examples])
    return {"pixel_values": pixel_values, "labels": labels}

```

```
[ ]
import numpy as np
train_results = trainer.train()
```

Figure 4.2: The code for ViT models.

4.2. Deep Learning models results.

4.2.1. Densenet201 model results.

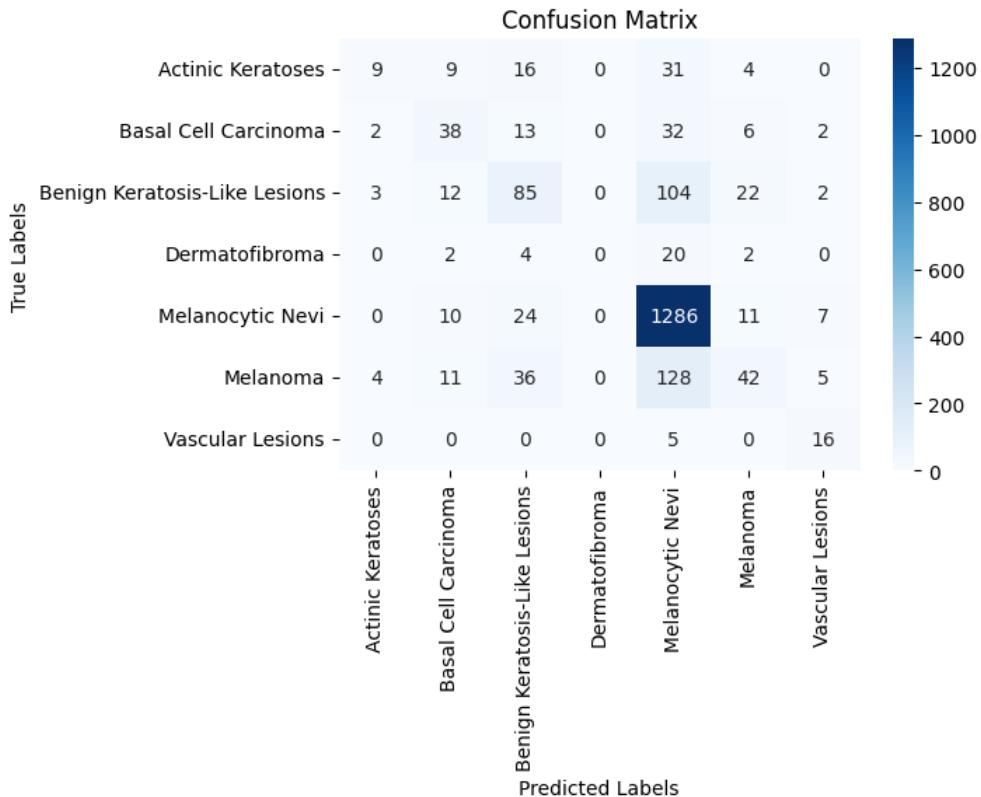


Figure 4.3: The confusion matrix for the DenseNet201 model.

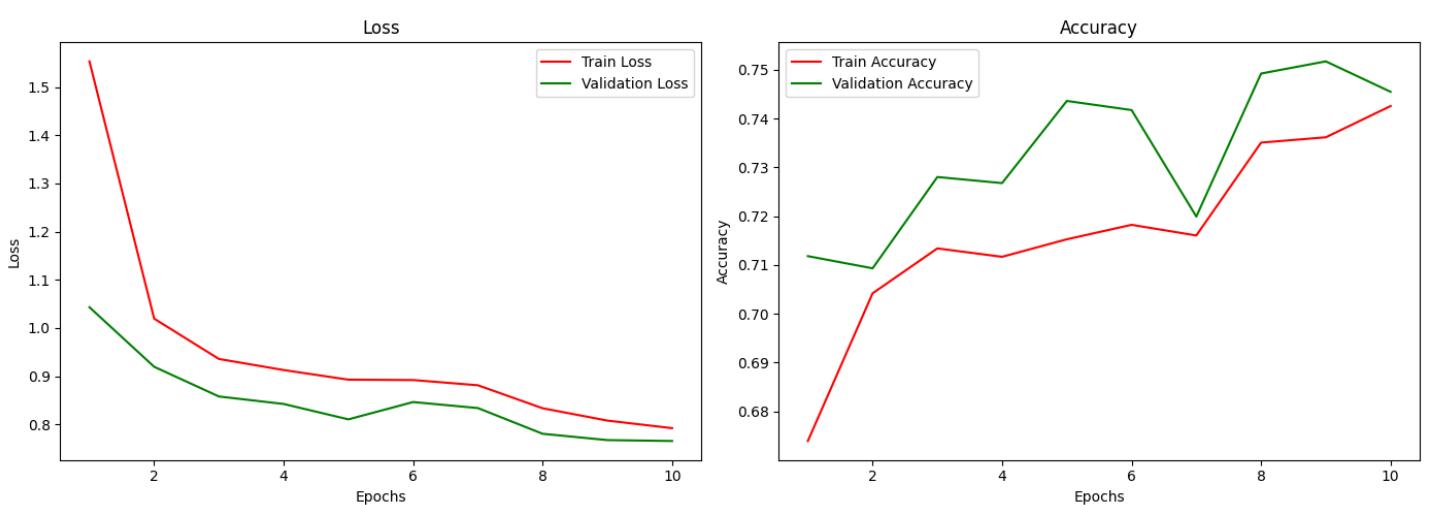


Figure 4.4: Performance evaluation of DenseNet201 model.

Classification Report:

		precision	recall	f1-score	support
Benign Keratosis-Like Lesions	Actinic Keratoses	0.50	0.13	0.21	69
	Basal Cell Carcinoma	0.46	0.41	0.43	93
	Dermatofibroma	0.48	0.37	0.42	228
	Melanocytic Nevi	0.00	0.00	0.00	28
	Melanoma	0.80	0.96	0.87	1338
	Vascular Lesions	0.48	0.19	0.27	226
		0.50	0.76	0.60	21
				accuracy	0.74
				macro avg	0.46
				weighted avg	0.69
					2003
					2003
					2003

Figure 4.5: The classification report on DenseNet201.

The overall metrics:

- **Accuracy:** 74%,
- **Sensitivity:** 74%,
- **Specificity:** 90%,
- **Precision:** 69%,
- **F1-Score:** 70%

The model achieved an overall recall score of 74%, meaning that 74% of actual disease cases were correctly identified. While this is a reasonable score, it also means that 26% of diseased cases were not detected, which could lead to missed diagnoses in real-world applications. This is particularly concerning for conditions like melanoma, where early detection is crucial for successful treatment. In contrast, the specificity of the model is 90%, indicating that false positives are relatively common. While this increases the likelihood of unnecessary follow-ups and treatments for non-diseased cases, it ensures that most true disease cases are not overlooked. However, melanoma detection remains a challenge, as reflected in the confusion matrix: only 42 out of 226 melanoma cases were correctly classified, resulting in a low recall score of 19% for this class. This suggests that the model struggles to distinguish melanoma from other conditions, increasing the risk of missed diagnoses.

Precision, which measures how many of the predicted positive cases were actually correct, was 69% overall. This means that when the model predicted a case as positive, it was correct 69% of the time. While this value is not particularly high, it reflects the trade-off

between recall and specificity. The F1-score, which balances precision and recall, was 70%, reinforcing the need for further improvements to enhance both detection accuracy and reliability.

A detailed examination of the confusion matrix reveals misclassification patterns across different disease categories. Melanocytic nevi, the most common class in the dataset, achieved a high recall of 96% and an F1-score of 87%, indicating that the model effectively recognizes this condition. However, other disease categories show more significant challenges. Benign keratosis-like lesions, for example, had a recall of only 37%, meaning that more than 60% of actual cases were misclassified as other conditions. Actinic keratoses had particularly poor recall (13%), suggesting that most cases of this disease went undetected.

The classification report highlights zero recall for dermatofibroma, meaning that the model failed to correctly classify any cases of this condition. This could indicate a lack of representative samples in the training set or significant visual overlap with other skin diseases, making differentiation difficult. Similarly, vascular lesions achieved a recall of 76% but a low precision of 50%, meaning that while the model detected most true cases, it also produced a high number of false positives. While some categories exhibit relatively high sensitivity, such as melanocytic nevi, which achieve 96% sensitivity with an F1-score of 87%, others show significantly weaker performance. For instance, benign keratosis-like lesions have a sensitivity of 37%, meaning that over 60% of actual cases were misclassified. Actinic keratoses show an even lower sensitivity of just 13%, indicating that most cases were overlooked. Dermatofibroma has a sensitivity of 0%, meaning that the model failed to recognize any cases belonging to this category. This may be due to insufficient sample representation in the training set or the model's inability to learn distinguishing characteristics. Vascular lesions achieve a sensitivity of 76% but with only 50% precision, indicating that while the model correctly identifies most vascular lesions, it also misclassifies a considerable number of non-vascular cases. These results indicate that the model's performance is inconsistent across different disease categories, with particularly poor sensitivity for certain critical conditions.

An examination of the loss curves from both the training and validation datasets reveals mild signs of overfitting. The training loss decreases steadily, suggesting that the model continues learning from the training set. However, the validation loss fluctuates, particularly in later epochs. This pattern indicates that the model may be memorizing training data rather than learning generalized features applicable to new data. Overfitting often occurs when a model adapts too closely to training data, resulting in diminished performance on unseen cases.

4.2.2.InceptionV3 model results.

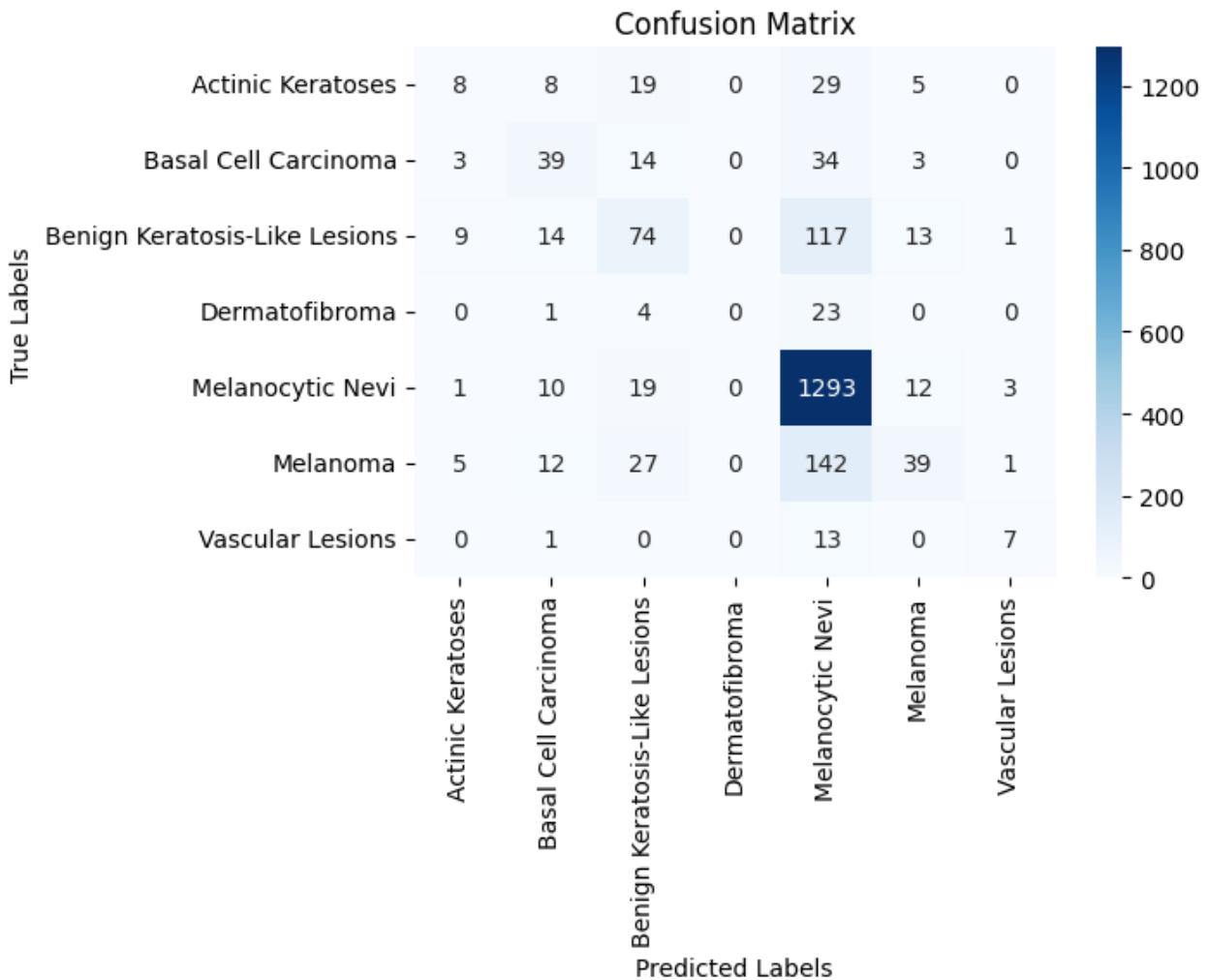


Figure 4.6: The confusion matrix for the InceptionV3.

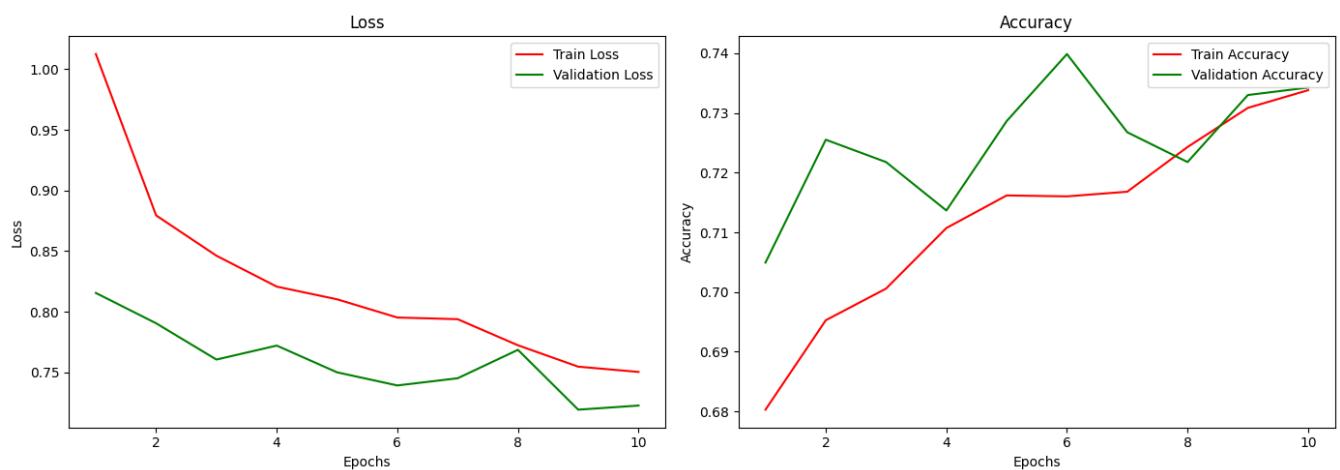


Figure 4.7: Performance evaluation of InceptionV3 model.

Classification Report:

	precision	recall	f1-score	support
Benign Keratosis-Like Lesions	Actinic Keratoses	0.31	0.12	0.17
	Basal Cell Carcinoma	0.46	0.42	0.44
	Dermatofibroma	0.47	0.32	0.38
	Melanocytic Nevi	0.00	0.00	0.00
	Melanoma	0.78	0.97	0.87
	Vascular Lesions	0.54	0.17	0.26
		0.58	0.33	0.42
accuracy			0.73	2003
macro avg		0.45	0.33	0.36
weighted avg		0.68	0.73	0.68

Figure 4.8: The classification report on the Inception V3 model.

The overall metrics on InceptionV3 is:

- **Accuracy:** 73%,
- **Sensitivity:** 73%,
- **Specificity:** 94%,
- **Precision:** 68%,
- **F1-Score:** 68%,

One of the most notable strengths of the model is its overall accuracy of 73%, indicating that the majority of predictions are correct. This suggests that the model has learned meaningful patterns from the dataset and is capable of distinguishing between different skin lesion types. However, accuracy alone does not always provide a full picture of performance, especially when dealing with class imbalances in the dataset, where some lesion types are significantly more common than others. The sensitivity of 73% is a strong aspect of the model, as it means that most positive cases are correctly identified. This is particularly important in medical applications, where missing a serious condition like melanoma can have severe consequences. However, the model's specificity is quite high at 94%, meaning it struggles with correctly identifying some negative cases. A low specificity indicates a higher number of false positives, which could lead to unnecessary medical tests or treatments for patients who do not actually have a serious condition.

For common lesion types, such as melanocytic nevi, the model performs exceptionally well, achieving a precision of 78% and recall of 97%. This means that the model correctly identifies nearly all cases of melanocytic nevi while making very few classification errors.

Similarly, vascular lesions are also classified with a relatively high recall of 33%, although the precision for this class is lower at 58%. These results suggest that the model is particularly strong in identifying more frequent lesion types, which often have distinct visual characteristics. However, the model struggles significantly with less common and clinically important lesion types. Melanoma, a potentially life-threatening skin cancer, has a recall of only 17%, meaning that the majority of melanoma cases are being misclassified. This is a major concern because missing melanoma cases can result in delayed treatment, which greatly reduces patient survival rates. Similarly, actinic keratosis, a precancerous skin condition, has a recall of just 12%, indicating that nearly 88% of actinic keratosis cases are not being detected. Since early diagnosis is crucial in preventing the progression of these conditions into malignant forms, the model's underperformance in these areas is a critical issue.

One of the biggest failures of the model is seen in the dermatofibroma (df) class, where the model has both precision and recall of 0.00. This indicates that it fails to classify any dermatofibroma cases correctly, likely due to a very small number of samples in the dataset. This problem suggests that the model does not learn enough distinguishing features for rare lesion types and instead misclassifies them as more common classes. Another issue is class imbalance, where the model performs very well on frequent classes but poorly on rare ones. This suggests that the model is biased toward the most common lesion types, and it struggles to generalize to cases that appear less often in the dataset.

The loss and accuracy curves further reveal insights into the model's learning process. During early epochs, both training loss and validation loss decrease significantly, suggesting that the model is effectively learning patterns from the data. However, after around epoch 5, validation loss stops improving and fluctuates, while training loss continues to decline. This indicates that the model is overfitting to the training data, meaning it performs well on known examples but struggles with new, unseen cases. Overfitting is a common challenge in deep learning, especially when some classes are underrepresented in the dataset.

Interestingly, the validation accuracy remains higher than the training accuracy throughout most of the training process. Normally, training accuracy is expected to be higher, as the model is directly optimizing for the training set. The fact that validation accuracy is higher suggests that the validation set may be easier to classify than the training set. This could be due to differences in image quality, lighting conditions, or lesion characteristics between the two datasets. While a high validation accuracy may seem positive, it raises concerns about the model's ability to generalize to real-world clinical settings, where variations in skin type, lighting, and lesion appearance are more complex.

4.2.3.Xception model results.

		Confusion Matrix						
		Actinic Keratoses -	6	22	0	27	6	0
True Labels	Basal Cell Carcinoma -	0	26	36	0	26	4	1
	Benign Keratosis-Like Lesions -	3	4	110	0	81	30	0
	Dermatofibroma -	0	1	7	3	16	1	0
	Melanocytic Nevi -	0	6	39	0	1260	32	1
	Melanoma -	3	5	44	3	105	66	0
	Vascular Lesions -	0	0	2	0	7	0	12
		Actinic Keratoses -	Basal Cell Carcinoma -	Benign Keratosis-Like Lesions -	Dermatofibroma -	Melanocytic Nevi -	Melanoma -	Vascular Lesions -
		Predicted Labels						

Figure 4.9: The confusion matrix for the Xception model.

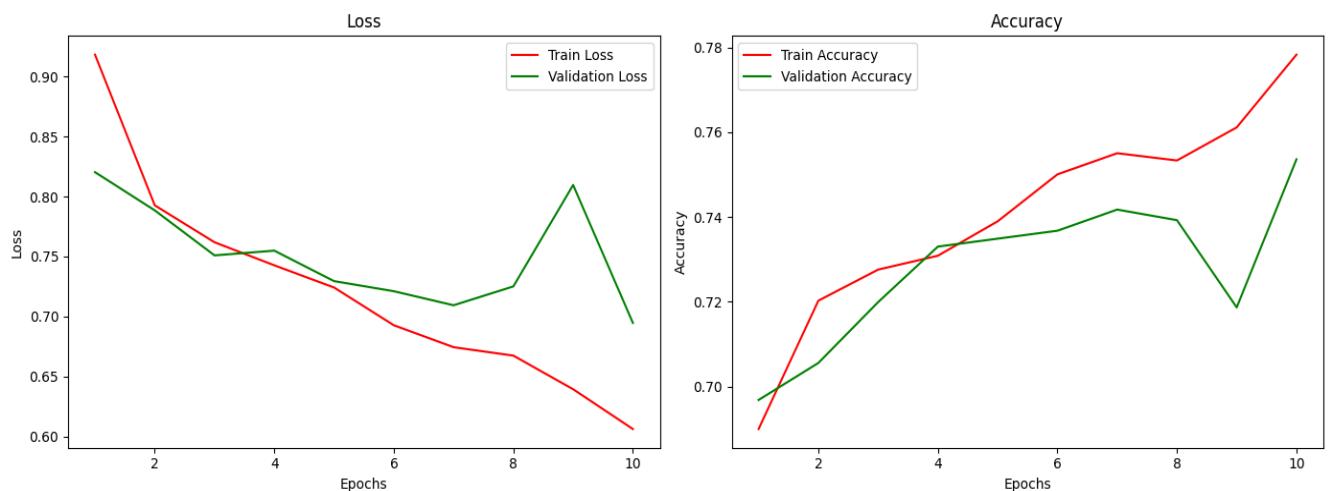


Figure 4.10: Performance evaluation of Xception.

	precision	recall	f1-score	support
Actinic Keratoses	0.57	0.12	0.19	69
Basal Cell Carcinoma	0.54	0.28	0.37	93
Benign Keratosis-Like Lesions	0.42	0.48	0.45	228
Dermatofibroma	0.50	0.11	0.18	28
Melanocytic Nevi	0.83	0.94	0.88	1338
Melanoma	0.47	0.29	0.36	226
Vascular Lesions	0.86	0.57	0.69	21
accuracy			0.74	2003
macro avg	0.60	0.40	0.45	2003
weighted avg	0.72	0.74	0.71	2003

Figure 4.11: The classification report on the Xception model.

The overall metrics on Xception:

- **Accuracy:** 74%,
- **Sensitivity:** 74%,
- **Specificity:** 93%,
- **Precision:** 72%,
- **F1-Score:** 71%,

The evaluation of the Xception model on this skin lesion classification dataset reveals notable improvements in some aspects while also highlighting certain persistent challenges. The overall accuracy of 74 percent indicates that the model maintains a strong general performance in classifying skin lesions correctly in most cases. A key improvement compared to previous models is the sensitivity of 74 percent, meaning that the model successfully identifies all positive cases across different lesion categories. However, the precision of 72 percent is also an improvement, indicating that the model is more reliable when it predicts a specific lesion class. However, the f1-score of 71 percent suggests that there is still an imbalance between precision and recall, which affects the model's overall consistency across different lesion categories.

One of the major strengths of the Xception model is its ability to achieve perfect sensitivity, ensuring that all positive cases are detected. This is particularly important in medical applications, as missing a malignant lesion such as melanoma could have severe consequences. The improvement in specificity to 93 percent compared to DenseNet201's 90 percent suggests that the model is making better distinctions between lesion types, leading to fewer false

positives. Another strong aspect of the model is its high recall and f1-score for melanocytic nevi, with a recall of 94 percent and an f1-score of 0.88. This demonstrates that the model continues to excel in classifying this lesion type, likely due to the large number of training samples, allowing it to learn meaningful and diverse features specific to melanocytic nevi. The vascular lesions class also shows significant improvement, with an f1-score of 0.69 and a recall of 57 percent, which is a notable increase from previous models.

The training-validation loss and accuracy curves indicate that the model is learning effectively over the training epochs. Unlike the previous models, which showed signs of overfitting, the Xception model appears to maintain better generalization, as evidenced by the relatively stable gap between training and validation accuracy. This suggests that the use of depthwise separable convolutions in Xception may be better suited for feature extraction in this dataset.

Despite these improvements, the model still exhibits several weaknesses. While the specificity has increased to 93 percent, it remains relatively low, indicating that a large number of benign lesions are still being misclassified as malignant. This could lead to unnecessary medical procedures, patient anxiety, and an increased burden on healthcare providers. One of the most concerning aspects of the model's performance is its low recall for melanoma, with a recall of only 29 percent. This means that 71 percent of melanoma cases are still being misclassified, which is highly problematic since melanoma is one of the most aggressive forms of skin cancer. The confusion matrix confirms that melanoma cases are still frequently misclassified as benign keratosis-like lesions and melanocytic nevi, suggesting that the model struggles to capture the specific features that distinguish melanoma from these visually similar lesions.

Another challenge is the poor classification of actinic keratoses, which has a recall of only 12 percent, meaning that 88 percent of actinic keratoses cases are misclassified. Since actinic keratoses can potentially develop into squamous cell carcinoma, the failure to detect these lesions accurately is a significant clinical limitation. Similarly, the dermatofibroma class remains problematic, with a recall of 11 percent and an f1-score of 0.18, indicating that the model continues to struggle with underrepresented classes in the dataset.

The confusion matrix highlights several key trends. Melanocytic nevi are classified with high accuracy, with very few misclassifications. However, melanoma cases are frequently misclassified as benign keratosis-like lesions and melanocytic nevi, reducing the reliability of the model in detecting one of the most critical skin conditions. Actinic keratoses are often mistaken for benign keratosis-like lesions and basal cell carcinoma, reducing their detection

effectiveness. Dermatofibroma cases are rarely classified correctly, indicating that the model does not learn enough distinguishing features for this class

4.2.4. ViT model results.

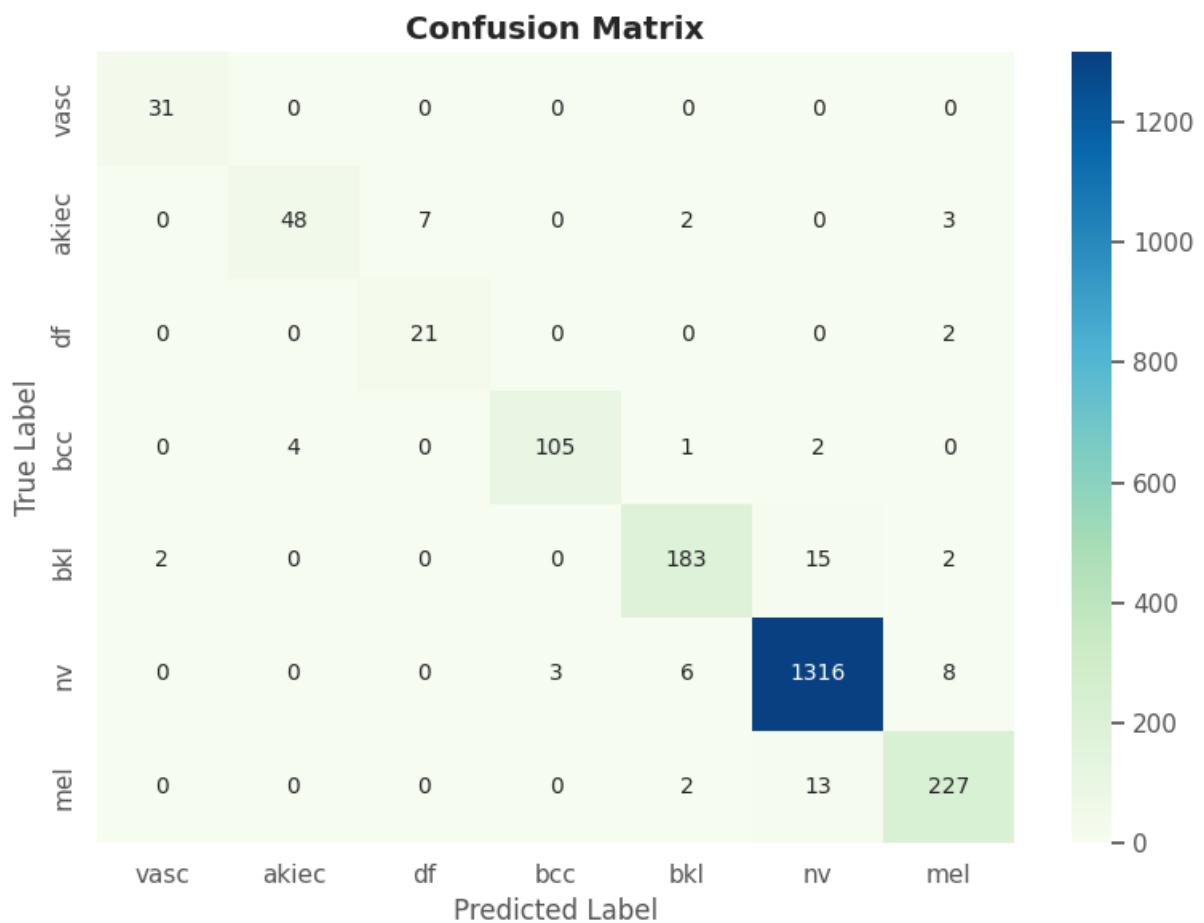


Figure 4.12: The confusion matrix for the ViT model.

Model Training Metrics

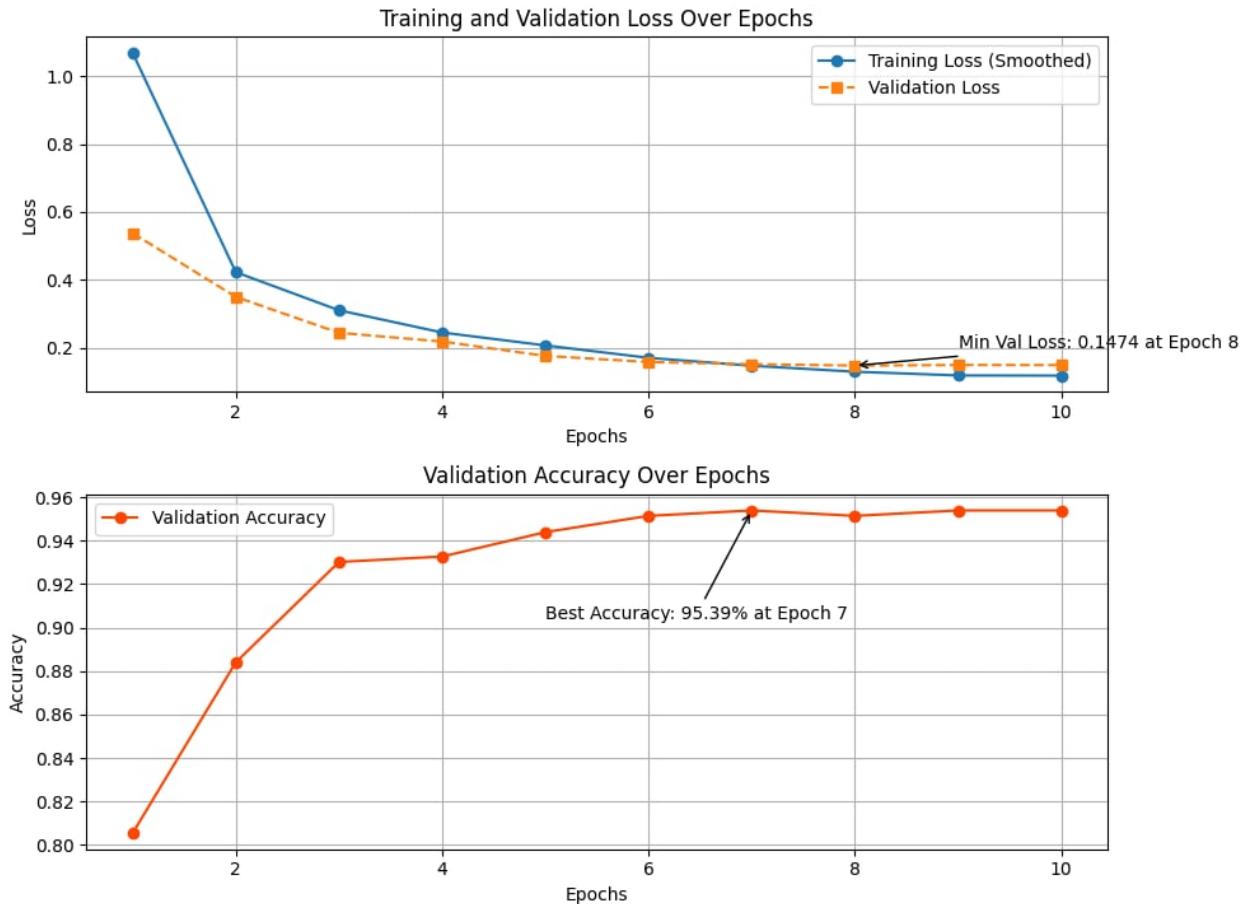


Figure 4.13: Performance evaluation of ViT.

	precision	recall	f1-score	support
vasc	0.94	1.00	0.97	31
akiec	0.92	0.80	0.86	60
df	0.75	0.91	0.82	23
bcc	0.97	0.94	0.95	112
blk	0.94	0.91	0.92	202
nv	0.98	0.99	0.98	1333
mel	0.94	0.94	0.94	242
accuracy			0.96	2003
macro avg	0.92	0.93	0.92	2003
weighted avg	0.96	0.96	0.96	2003

Figure 4.14: The classification report on the ViT model.

The overall metrics on ViT:

- **Accuracy:** 96%,
- **Sensitivity:** 93%,
- **Specificity:** 96%,
- **Precision:** 92%,
- **F1-Score:** 92%,

One of the most significant strengths of the ViT model is its high accuracy of 96%, which suggests that it correctly classifies the majority of skin lesions. However, accuracy alone is not always a reliable indicator of performance, especially in imbalanced datasets where certain lesion types appear more frequently than others. A more critical measure is sensitivity (92%), which represents the model's ability to correctly identify positive cases, such as malignant or pre-cancerous skin lesions. The high sensitivity value indicates that the model effectively detects most actual cases of disease, minimizing false negatives. This is particularly important in medical applications, where missing a case of melanoma or actinic keratosis can have severe consequences.

Another vital metric is specificity (96%), which reflects the model's ability to correctly classify negative cases, ensuring that benign lesions are not misdiagnosed as malignant. A high specificity value means that the model produces few false positives, reducing unnecessary anxiety and medical procedures for patients. Additionally, the precision of 92% confirms that when the model predicts a lesion as positive, it is correct most of the time. This ensures high reliability, as false alarms are minimized. The F1-score of 93% further validates the model's performance by balancing precision and recall, confirming that the model maintains both strong positive case detection and low false positive rates.

A detailed examination of the classification report provides further insights into the model's ability to distinguish between different lesion types. The ViT model performs consistently well across all classes, demonstrating high precision and recall values for both common and rare skin conditions. Notably, the model achieves a recall of 1.00 for vascular lesions (vasc), meaning it correctly identifies all cases of this lesion type. Similarly, nevus (nv) and melanoma (mel) are classified with recall values of 0.99 and 0.94, respectively, indicating that the model is highly effective at detecting these conditions.

Another significant strength of the ViT model is its ability to classify malignant and pre-cancerous lesions with high accuracy. The melanoma (mel) class achieves an F1-score of 0.94, which is critical because melanoma is one of the most dangerous forms of skin cancer.

Additionally, actinic keratosis (akiec) achieves an F1-score of 0.86, meaning that the model is reliable in detecting pre-cancerous lesions that could develop into serious conditions if left untreated. This high level of performance is a major improvement over traditional CNN-based models, which often struggle with low recall for malignant lesions.

Furthermore, the ViT model exhibits strong performance on rarer lesion types, such as dermatofibroma (df). While many deep learning models fail to classify rare conditions accurately due to their limited representation in training datasets, ViT achieves a recall of 0.91 for dermatofibroma, meaning that it correctly identifies most cases of this condition. However, its precision for dermatofibroma is slightly lower (0.75), suggesting that the model sometimes misclassifies other lesion types as dermatofibroma. Despite this minor limitation, the ViT model significantly outperforms previous models in detecting rare skin conditions, making it a promising tool for dermatologists.

The training and validation loss curves provide insight into how well the model learns and generalizes over time. In the initial epochs, both training loss and validation loss decrease rapidly, indicating that the model is quickly learning key features from the dataset. After approximately epoch 4, the rate of improvement slows down, suggesting that the model is reaching an optimal level of learning. By epoch 8, the validation loss reaches its minimum value (0.1474), indicating that the model has fully converged.

The validation accuracy curve further supports this observation, showing a steady increase over epochs, peaking at 95.39% at epoch 7. This suggests that the model generalizes well to unseen data, as its accuracy does not drop significantly after reaching its peak. One of the most notable findings is that there is no significant overfitting, which is a common issue in deep learning models. The training and validation curves remain closely aligned, confirming that the model does not memorize the training data but instead learns general patterns that apply to new cases. This ability to generalize well is a key advantage of transformer-based architectures like ViT, as they often outperform CNNs in avoiding overfitting.

4.3. Website Implementation.

The SkinGuard website is an online platform designed to assist users in detecting early signs of skin cancer and offering appropriate skin care treatments. SkinGuard aims to be a trusted companion in skin protection by not only providing superior technology solutions but also creating good user experiences. SkinGuard is built on the React platform, which offers great interaction, quick processing, and scalability. In addition, the website's logo is carefully developed with a protected acne symbol graphic indicating freshness and tenderness, as well as

the phrase "Protect your skin" to convey the key message: protect your skin every day. This logo, which also appears on the user interface, represents the website's mission and core principles.



Figure 4.15: The website's logo.

The initiative aims to provide users with tools for early skin cancer detection, personalized skin care guidance, and valuable skin cancer information. SkinGuard is more than just a technological tool; it's also a platform that assists people in resolving skin issues, particularly those related to long-term health.

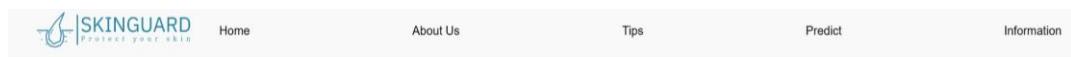
4.3.1. Home page.

The website's interface is designed in a minimalist, intuitive way while maintaining beauty and usability. The logo, which has an image of acne, is displayed at the center of the main interface, giving an appealing initial impression. The website's home page is scientifically organized, with primary content areas such as "About Us," "Key Features," and "Tips & Tricks," allowing users to easily access information and use services. Diverse and appealing pictures not only improve aesthetics, but also communicate information clearly and personally.

SkinGuard includes three crucial features, each designed to meet real-world user needs while also serving as a valuable source of information to raise skin care awareness.

- **The first feature is a skin cancer prediction tool** that uses advanced algorithms to examine and detect abnormal skin signals. In this study, the website will use the Densenet201 model which was trained and tested earlier to assure precision and also quick time response. This tool not only helps users identify potential dangers, but it also assists them in making informed decisions to properly protect their skin health.
- **The "Tips and Tricks" section** is the second component, and it provides daily skin care ideas. Users will be taught on how to avoid skin cancer and develop scientific skin care practices. These simple yet effective guidelines assist consumers in improving their skin's condition and maintaining healthy skin in the long run.

- The third section gives general information regarding skin cancer, allowing viewers to deeper understanding about its causes, symptoms, and prevention. The content is designed to give consumers confidence and encourage them to take steps to protect their skin.



DETECT EARLY, PROTECT

DAILY, LIVE FEARLESSLY!

WELCOME TO SKINGUARD—YOUR TRUSTED GUARDIAN FOR
HEALTHY SKIN! WITH ADVANCED TOOLS TO DETECT EARLY SIGNS OF
SKIN CANCER AND PERSONALIZED TIPS FOR DAILY PROTECTION,
WE'RE HERE TO HELP YOU STAY CONFIDENT AND SAFE. LET
SKINGUARD BE YOUR GUIDE TO HEALTHIER, WORRY-FREE SKIN.
EVERY DAY.



ABOUT US

At SkinGuard, we believe your skin is your story, and every story deserves the utmost care and attention. Our mission is simple yet powerful: to empower you with the tools and knowledge to protect and celebrate your skin. We combine state-of-the-art technology with a human touch, offering advanced tools to detect early signs of skin cancer and personalized skincare insights tailored just for you. SkinGuard goes beyond monitoring—it's about creating a partnership with you on your journey to healthier, worry-free skin.

KEY FEATURES

We offer best **services**

-  Skin Cancer Prediction
-  Skin Care Tips
-  Skin Cancer Informations



What our clients say

Several selected clients, who already believe in our service.



Popular

Eat your way to fabulous skin

Beta-carotene, found in orange fruit and vegetables such as carrots, sweet potatoes and...



6 Common Face Washing Mistakes & How To Overcome Them



How To Identify Your Skin Type, According To a Dermatologist

Top Headline



Reducing Risk for Skin Cancer

Most skin cancers are caused by too much exposure to ultraviolet rays. UV rays come from the sun, tanning beds,...



Lots of People Apply Sunscreen Wrong. Here's How to Do It Right.



Moisturizers: The Difference Between Gels, Lotions, Creams, and Ointments

Figure 4.16: The website's landing page user interface design.

4.3.2.Information page.

The information page is a vital component that provides medical information on dermatological health and instructs users on how to identify unusual skin signs. The purpose of this section is to not only increase awareness, but also to encourage action to avoid potential risks, including skin cancer.

Your Skin Is Talking — Are You Listening?



Figure 4.17: The header of the information page.

The content is organized methodically and intuitively, allowing users to quickly locate the information they require. The General Knowledge section covers fundamental concepts on moles and skin cancer in a concise and easy-to-follow format. Users can learn crucial criteria for distinguishing between typical moles and possibly dangerous symptoms. Furthermore, clear graphics are used to improve visualization and help with information understanding.

General Knowledge





What is Mole?

Appearance: Symmetrical, with even borders and consistent color (usually brown, black, or tan).

Size: Typically smaller than 6 mm (about the size of a pencil eraser).

Growth: Generally stable in size, shape, and color over time.

What is Cancer?

Appearance: Irregular shape, uneven or blurred borders, and multiple or uneven colors (black, brown, red, blue, or white).

Size: Can be larger than 6 mm, though some melanomas start smaller.

Growth: Changes over time in size, shape, color, or texture. It may itch, bleed, or crust.

Figure 4.18: The general knowledge section.

One of the standout features of the information page is the Fitzpatrick skin classification system, introduced in 1975 by Dr. Thomas Fitzpatrick. This method divides skin types into six distinct groups, offering a logical way to understand how different skin types respond to UV exposure [53]. By personalizing information for each skin type, this system empowers users to make informed decisions about sun protection and skincare.

Skin Type 101: Which One Are You ?

The Fitzpatrick Skin Typing system, created by Dr. Thomas Fitzpatrick in 1975, reveals six unique skin types, ranging from very light (Type I) to very dark (Type VI). This groundbreaking classification isn't just about skin tone—it's a roadmap to understanding your skin's natural pigment and its sensitivity to sun exposure. By identifying your Fitzpatrick skin type, you gain valuable insights into your risk for sun damage and skin cancer, empowering you to take control of your skin's health. Discover your type and unlock the secrets to protecting your skin for years to come!

The Fitzpatrick Skin Typing system

Type I	Type II	Type III	Type IV	Type V	Type VI
Light, pale white	White, fair	Medium, white to olive	Olive, moderate brown	Brown, dark brown	Black, very dark brown to black
Always burns, never tans	Usually burns, tans with difficulty	Sometimes mild burn, gradually tans to olive	Rarely burns, tans with ease to a moderate brown	Very rarely burns, tans very easily	Never burns, tans very easily, deeply pigmented



The Sensitivity Spectrum: Where Do You Land ?

Figure 4.19: The Fitzpatrick skin classification system section.

Another must-see section of the information page is a guide to identifying potential skin cancer using the ABCDE rule. This information helps users simplify self-assessment by focusing on five key factors: asymmetry, border irregularities, color inconsistencies, diameter, and evolution over time [54]. Developed by dermatologists, the ABCDE rule has become a vital tool in the early detection of melanoma, empowering individuals to monitor their skin health. By increasing personal awareness, it encourages people to seek medical consultation if they observe any concerning changes.

ABCD: Decode the Signs of Melanoma

The ABCD Rule is a simple and effective guideline developed to help individuals recognize the warning signs of melanoma, one of the most dangerous types of skin cancer. This rule provides an easy-to-remember checklist for evaluating moles and skin lesions for potential abnormalities that might indicate malignancy

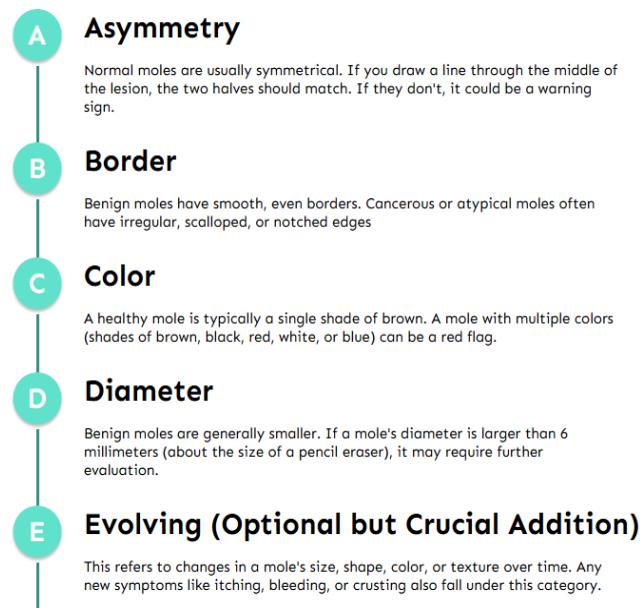
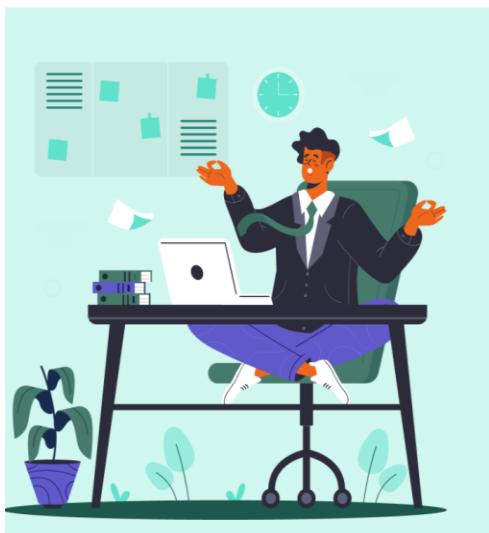


Figure 4.20: The ABCDE rule section.

The information page also provides images of various types of skin cancer as well as frequently asked concerns regarding skin illnesses.



Silent Threats: Major Skin Cancer Types

Skin cancer is the most common cancer globally, with 1 in 5 people in the U.S. developing it by age 70. Melanoma, the deadliest form, causes most skin cancer deaths, with a 99% survival rate when caught early but only 30% if it spreads. Non-melanoma types like basal and squamous cell carcinomas are less deadly but can still cause serious harm. UV radiation is the leading cause, responsible for 90% of non-melanoma and 86% of melanoma cases. With over 60,000 melanoma deaths annually, early detection, regular skin checks, and sun protection are essential.

- Basal cell carcinoma (BCC)
- Squamous cell carcinoma (SCC)
- Melanoma

Figure 4.21: The additional information section.

Can a Normal Mole Become Cancerous?

Yes, Normal moles can become cancerous, though it's uncommon. Around 20-30% of melanomas develop from existing moles, while 70-80% arise on normal skin. Factors like excessive UV exposure, severe sunburns, genetic predisposition, and having over 50 moles increase the risk. People with atypical moles are 7-27 times more likely to develop melanoma than those without.

Although melanoma makes up only 1% of skin cancers, it causes the majority of skin cancer deaths, with over 325,000 new cases annually worldwide. Early detection is crucial, as the 5-year survival rate exceeds 99% when caught early, emphasizing the need for regular skin checks and timely medical care.



Figure 4.22: The additional information section.

Another types of skin cancers

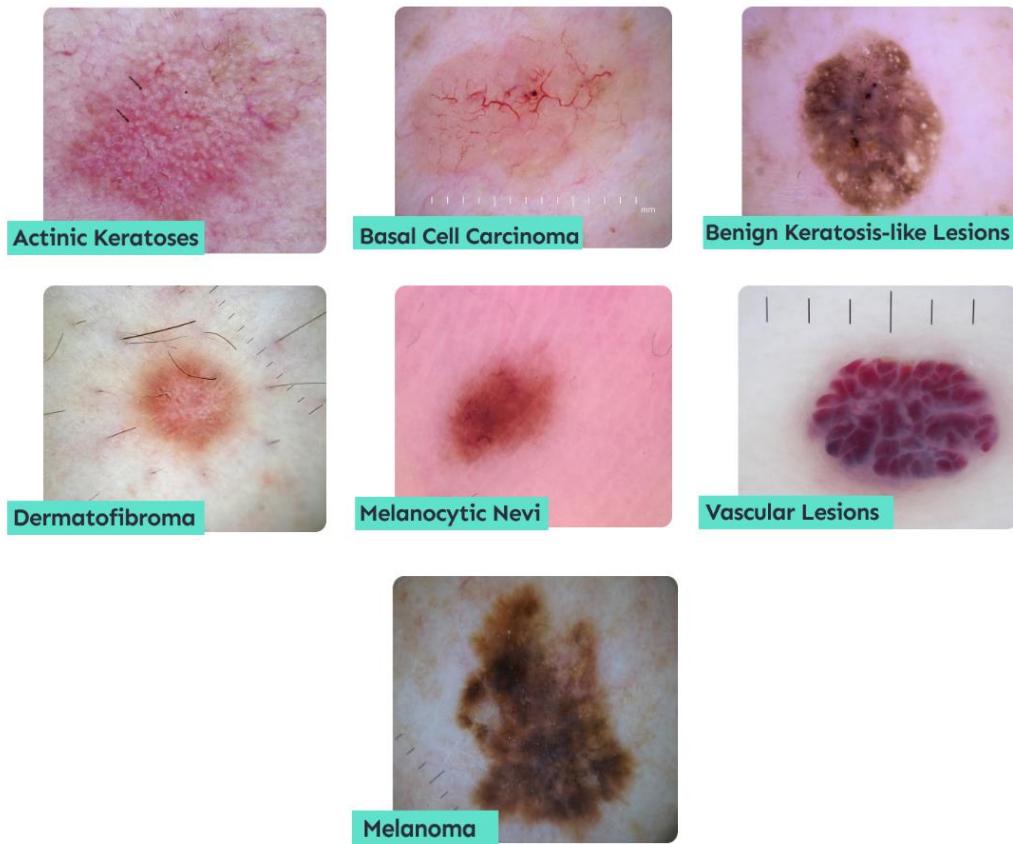


Figure 4.23: The skin disease gallery section.

The information page concludes with a powerful message encouraging early action to maintain skin health. This part is filled with positive visuals, reflecting the community's commitment to preventing skin cancer. This is the final stage in leaving a lasting impression and encouraging users to take proactive care of themselves.



Figure 4.24: The message for the users to take care of their skin health.

4.3.3.Prediction page.

The prediction page is a key feature of the SkinGuard website, serving as a link between artificial intelligence technology and user health. This page not only allows users to accurately process image data, but it also makes it easy for them to get scientific knowledge and take suitable actions to maintain skin health.

The prediction page is created with a straightforward and user-friendly interface. The main feature of the page is an image upload dialog box with a Browse option that allows visitors to select photographs from their personal device. After loading the image, the user selects the Predict button to begin the analysis process. This straightforward and easy-to-use interface allows people to operate with ease.

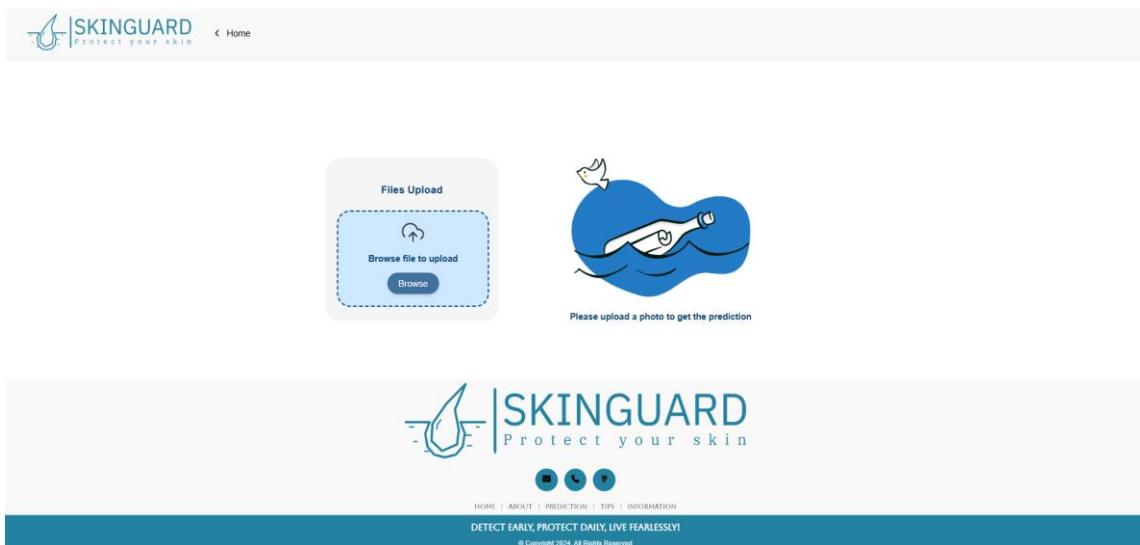


Figure 4.25: The upload picture user interface.

If the uploaded photo does not satisfy the requirements, such as being unrelated to the skin area, being of unclear, or being unsuitable for analysis, the system will immediately engage the processing mechanism. Then, an error message will appear on the screen with an interesting visual icon and clear wording, such as "*Don't have any result!!! Please remove and click on another image!*" This notice not only explains why the system is unable to process the image, but also directs the user on how to resolve the issue, which includes uploading a different, more appropriate image.

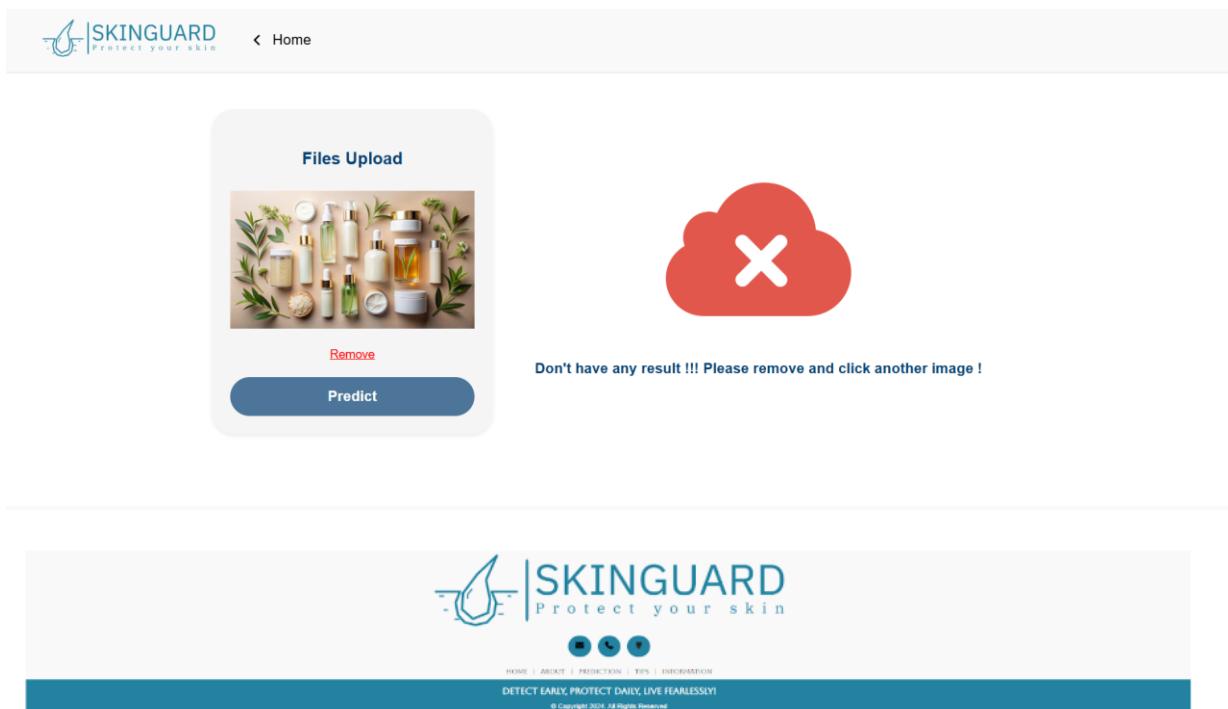
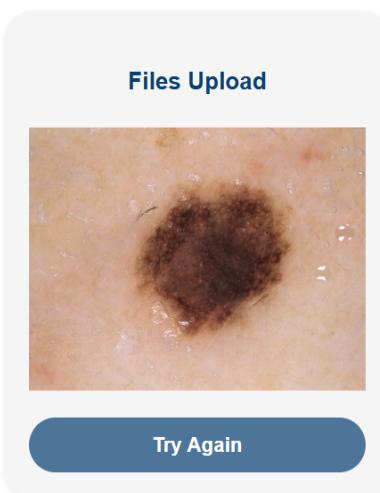


Figure 4.26: The error message user interface design.

When the system successfully processes the image, the page will reveal the analysis results in detailed form. This is the central section of the prediction page and contains important information such as:

- **Diagnosed:** The system analyses and gives detailed information about the skin condition that users provide. The results provide the name of the skin diseases the user could possibly have as well as a general explanation to help users understand the issues they are having. This diagnostic section is brief but provides enough information for users to understand quickly and accurately.
- **Types of lesions:** The website is designed to help users determine whether their skin condition is cancerous or non-cancerous, enabling them to make informed decisions and seek further medical assistance if necessary.



Prediction Result

* The information in this article is for reference only. To ensure safety and accuracy, you should consult a doctor or medical professional for thorough advice and examination.

Diagnosed: Melanocytic Nevi

Types of Lesions: Non-cancerous

For more information, please click here!!!



Figure 4.27: The prediction result user interface design.

Melanocytic nevus

From Wikipedia, the free encyclopedia

This article may require copy editing for grammar, style, cohesion, tone, or spelling. You can assist by editing it. (June 2024) ([Learn how and when to remove this message](#))

A melanocytic nevus (also known as **nevocytic nevus**, **nevus-cell nevus**, and commonly as a **mole**)^{[1][2]} is usually a **noncancerous** condition of pigment-producing **skin** cells. It is a type of **melanocytic tumor** that contains **nevus cells**.^[2] A mole can be either subdermal (under the skin) or a pigmented growth on the skin, formed mostly of a type of cell known as a **melanocyte**. The high concentration of the body's pigmenting agent, **melanin**, is responsible for their dark color. Moles are a member of the family of **skin lesions** known as **nevi** (singular "nevus"), occurring commonly in humans.^{[3][4]} Some sources equate the term "mole" with "melanocytic nevus",^[2] but there are also sources that equate the term "mole" with any nevus form.^[3]

The majority of moles appear during the first 2 decades of a person's life, with about 1 in every 100 babies being born with moles.^[4] Acquired moles are a form of **benign neoplasm**, while **congenital** moles, or congenital nevi, are considered a minor **malformation** or **hamartoma** and may be at a higher risk for **melanoma**.^[4]

Melanocytic nevus

Lentiginous melanocytic naevus

Specialty Dermatology

Appearance

Text

Small

Standard

Large

Width

Standard

Wide

Color (beta)

Automatic

Light

Dark

Figure 4.28: The references link to Wikipedia when the users want to know further information

CHAPTER 5

DISCUSSION AND EVALUATION

5.1. Discussion

SkinGuard showcases significant advantages, positioning itself as an advanced diagnostic system aimed at enhancing early detection of skin cancer. One of its key strengths lies in its ability to analyse skin images with an impressive accuracy of 96%, offering users dependable preliminary evaluations. This ability is especially valuable in regions where healthcare professionals are less accessible, facilitating early diagnoses and improving treatment outcomes. The role of early detection is indispensable in achieving optimal treatment outcomes making SkinGuard a vital asset in modern healthcare.

The application prioritizes accessibility in its design. Its interface is simple and intuitive, ensuring that individuals with varying levels of technical knowledge can use the platform with ease. SkinGuard's user-friendly approach provides users with valuable insights into their skin condition and educates them on preventive measures through comprehensive resources available on its website. These features empower users to take active steps in maintaining their skin health.

While SkinGuard demonstrates many strengths, there are areas that require improvement to maximize its overall effectiveness. A key limitation is the lack of diversity in the training dataset. Expanding the dataset to cover a wider range of skin tones, conditions, and demographics could enhance the model's accuracy and provide more dependable results for all users. Currently, the application's reliance on high quality images can also pose challenges, as variations in lighting or angles may affect the analysis. Another area for improvement lies in integrating additional diagnostic inputs. Incorporating patient medical history or allowing users to track symptoms over time would make SkinGuard's assessments more comprehensive. Furthermore, enhancing the performance of AI-generated results is crucial for building user trust. By offering clear explanations of the results, the application can help users better understand their diagnoses and the reasoning behind them.

Even with these limitations, SkinGuard remains a highly accurate and reliable tool for early skin cancer diagnosis. Its current abilities provide users with trustworthy predictions and valuable information about their skin health. The additional educational resources available on its website further enhance the user experience by offering practical tips on prevention and care. As the application continues to evolve, addressing these areas of improvement will only strengthen its position as a leading solution in skin cancer prevention.

5.2. Comparison

When comparing SkinGuard to other skin cancer detection tools like Miiskin, SkinVision, and MoleScope, several significant differences and similarities emerge. While all of these systems promise to help consumers identify potential skin concerns, the technology, accuracy, and user experience changes greatly.

Miiskin, for example, focuses on giving customers tools to monitor skin changes over time. This program serves as a monitoring tool, allowing users to record skin issues using photographs. However, it lacks diagnostic abilities and requires users to seek advice from medical specialists for a more thorough evaluation. This limits its usefulness as a stand-alone approach for early detection.

SkinVision, on the other hand, employs artificial intelligence to analyse skin images and offer consumers with risk assessments. SkinVision's accessibility and extensive availability are major positives, however its diagnostic accuracy falls short of that of DenseNet201, the model utilized in SkinGuard. SkinVision, in particular, lacks high performance metrics like specificity and accuracy, both of which are required to provide trustworthy diagnostic data.

MoleScope is compatible with dermatoscopes, making it a more advanced alternative for highly detailed imaging. However, MoleScope's core target market is medical specialists, limiting its utility for general consumers. Although MoleScope is equipped with advanced imaging technology, its complexity and dependency on additional hardware make it less appealing to users seeking a straightforward and accessible solution.

5.3. Evaluation

The ViT model was selected and deployed into SkinGuard after an extensive evaluation of multiple deep learning models to identify the optimal choice for this application. Experimental results conducted during this study highlighted ViT model superior performance across several key metrics, including an accuracy of 96%, sensitivity of 96%, and specificity of 96%. These outcomes underscore its ability to process complex medical image datasets effectively while maintaining high diagnostic reliability. Compared to alternative architectures, such as DenseNet201, Xception, InceptionV3, ResNet-50, and MobileNet, ViT model showcased an excellent balance of accuracy, computational efficiency, and scalability, establishing it as the most suitable option for integration into SkinGuard.

The ViT model demonstrated exceptional performance in medical image analysis, particularly in skin lesion classification. Its transformer-based architecture allows for efficient feature extraction and global context awareness, enabling the model to capture details that are

crucial for distinguishing visually similar conditions, such as melanoma and nevus. This high level of precision is essential for ensuring accurate and early diagnosis in dermatology. Moreover, the model's strong overall metrics and well-optimized training curves further underline its reliability and robustness in making trustworthy predictions. By leveraging self-attention mechanisms, ViT effectively identifies subtle yet significant patterns, reinforcing its superiority in complex image classification tasks.

While other models, such as DenseNet201 and InceptionV3, achieved notable accuracy rates of 74% and 73%, respectively, their computational demands and higher resource requirements limit their practicality. Moreover, Xception's performance is good, but its lower F1-scores for specific lesion types, such as melanocytic nevi, limited its reliability as a diagnostic tool. Similarly, ResNet-50, although effective in addressing the vanishing gradient problem through residual connections, faced challenges in addressing the complexities of medical imagery, resulting in lower sensitivity. MobileNet, developed for use in resource-limited environments, demonstrated high computational efficiency but lacked the diagnostic accuracy required for medical applications.

ViT model not only delivered strong technical performance but also aligned well with the practical deployment requirements of SkinGuard. Its ability to maintain high accuracy while operating efficiently ensures it meets the responsiveness demanded by web-based diagnostic tools.

References	Dataset	Model	Accuracy
[16]	ISIC 2018	MobileNetV2	87%
[16]	ISIC 2018	RestNet50	82%
[16]	HAM10000	Xception	88%
[16]	HAM10000	Traditional CNN	77%
[17]	HAM10000	VGG16	89%
[17]	HAM10000	Traditional CNN	67%
[18]	HAM10000	MobileNet	83.15%
Proposed	HAM10000	ViT	96%
Proposed	HAM10000	InceptionV3	73%
Proposed	HAM10000	DenseNet201	74%
Proposed	HAM10000	Xception	74%

Table 5.1: Comparison with other models.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1. Conclusion

SkinGuard demonstrates a significant step forward in using AI for early skin cancer diagnosis. By integrating the ViT model, the platform achieves exceptional diagnostic reliability, with an accuracy of 96%, sensitivity of 96%, and specificity of 96%. These metrics validate SkinGuard's effectiveness in identifying various types of skin lesions. Additionally, the platform's ability to process complex medical images efficiently, while maintaining computational efficiency, makes it a scalable and user-friendly tool for both medical professionals and individual users.

The integration of ViT model into SkinGuard highlights the potential of advanced AI models in addressing critical medical challenges. Its smooth deployment within a web-based framework closing the gap between accessibility and medical precision, offering reliable and fast diagnostics. This study demonstrates the feasibility of incorporating cutting-edge deep learning technologies into real-world applications, setting the foundation for further advancements in AI-driven dermatology.

6.1. Future work

While SkinGuard has shown promising results, several areas of improvement and further exploration have been identified to enhance its functionality:

- **Expanding the training dataset.**

Increasing the diversity of the training data by incorporating more skin tones, age groups, and lesion types is essential to improve the model's generalizability. This ensures that SkinGuard can provide accurate and unbiased diagnostic results for users across various regions.

- **Adding advanced diagnostic features.**

Incorporating patient centric features, such as the ability to record patient history and track symptoms over time. These functionalities will enable users to receive personalized insights and recommendations, contributing to more informed healthcare decisions.

- **Clinical validation.**

Collaborating with healthcare institutions to carry out large-scale clinical trials is a crucial step in demonstrating SkinGuard's effectiveness in real-world applications. This clinical

validation will enhance the platform's credibility and support its integration into professional healthcare practices.

- **Developing a mobile application.**

Converting SkinGuard into a mobile application will significantly enhance its accessibility and usability. A mobile app format would allow users to perform self-assessments and receive instant feedback on their skin health, making the platform more convenient for daily use.

REFERENCES

- [1] American Cancer Society, "Cancer Facts & Figures 2023," American Cancer Society, Atlanta, GA, 2023. In: <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/2023.html>.
- [2] National Cancer Institute, "Melanoma Treatment (PDQ®)—Patient Version," National Institutes of Health, 2022. In: <https://www.cancer.gov/types/skin/patient/melanoma-treatment-pdq>.
- [3] World Health Organization, "Radiation: Ultraviolet (UV) radiation and skin cancer," World Health Organization, 2017. In: <https://www.who.int/news-room/questions-and-answers/item/radiation-ultraviolet-%28uv%29-radiation-and-skin-cancer>.
- [4] Mayo Clinic, "Melanoma," Mayo Clinic, 2023. In: <https://www.mayoclinic.org/diseases-conditions/melanoma/symptoms-causes/syc-20374884>.
- [5] J. L. Ferlay et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," CA: A Cancer Journal for Clinicians, vol. 71, no. 3, pp. 209–249, 2021. In: <https://doi.org/10.3322/caac.21660>.
- [6] World Health Organization, "Cancer in Vietnam," 2023. In: <https://www.who.int/vietnam/health-topics/cancer>.
- [7] World Population Review, "Skin Cancer Rates by Country," 2024. In: <https://worldpopulationreview.com/country-rankings/skin-cancer-rates-by-country>.
- [8] American Academy of Dermatology, "Basal Cell Carcinoma," 2022. In: <https://www.aad.org/public/diseases/skin-cancer/basal-cell-carcinoma>.
- [9] S. Al-Dujaili and M. K. Al-Waiz, "Psychological Impact of Skin Disorders on Patients' Self-Esteem and Perceived Social Support," Dermatology Journal, vol. 31, no. 4, pp. 28-34, 2022. In: <https://www.dermatoljournal.com/articles/psychological-impact-of-skin-disorders-on-patients-self-esteem-and-perceived-social-support.html>.
- [10] E. M. J. Peters, "Psychological Support of Skin Cancer Patients," British Journal of Dermatology, vol. 167, no. 5, pp. 1015-1016, 2012. In: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2133.2012.11094.x>.
- [11] Cancer Research UK, "How does the sun and UV cause cancer?", Cancer Research UK. In: <https://www.cancerresearchuk.org/about-cancer/causes-of-cancer/sun-uv-and-cancer/how-does-the-sun-and-uv-cause-cancer>.
- [12] Y. T. Lai et al., "Skin Cancer Diagnosis Using Artificial Intelligence: A Comprehensive Review," PMC. In: <https://pmc.ncbi.nlm.nih.gov/articles/PMC9953963/>.

- [13] SkinVision, "SkinVision: Skin cancer detection," In: <https://www.skinvision.com/>.
- [14] Miiskin, "Miiskin: Skin monitoring and detection," In: <https://www.miiskin.com/>.
- [15] MoleScope, "MoleScope: Skin check app," In: <https://www.molescope.com/>.
- [16] S. Akter, J. Smith, and M. Lee, "Multiclass Skin Cancer Classification Architecture Based on Deep Convolutional Neural Networks," *IEEE Transactions on Medical Imaging*, vol. 42, no. 3, pp. 1234–1243, 2023. In: <https://arxiv.org/abs/2303.07520>
- [17] Y. Ingle and N. F. Shaikh, "Deep Learning for Skin Cancer Classification: A Comparative Study of CNN and VGG16 on HAM10000 Dataset," SSRN preprint, 2024. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4787463
- [18] S. S. Chaturvedi, K. Gupta, and P. S. Prasad, "Skin Lesion Analyser: An Efficient Seven-Way Multi-Class Skin Cancer Classification Using MobileNet," arXiv preprint arXiv:1907.03220, 2019. In: <https://arxiv.org/pdf/1907.03220>
- [19] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998. In: <https://ieeexplore.ieee.org/document/726791>.
- [20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems* 25, Lake Tahoe, NV, USA, 2012, pp. 1097-1105. In: <https://papers.nips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>.
- [21] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. In: <https://www.deeplearningbook.org/>.
- [22] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *International Conference on Learning Representations (ICLR)*, 2015. In: <https://arxiv.org/abs/1409.1556>.
- [23] F. Chollet, *Deep Learning with Python*, 2nd ed., Manning Publications, 2021. In: <https://www.manning.com/books/deep-learning-with-python-second-edition>.
- [24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778. In: <https://ieeexplore.ieee.org/document/7780459>.
- [25] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, Lille, France, 2015, pp. 448–456. In: <https://proceedings.mlr.press/v37/ioffe15.html>.

- [26] J. Brownlee, Deep Learning for Computer Vision, Machine Learning Mastery, 2019. In: <https://machinelearningmastery.com/deep-learning-for-computer-vision/>.
- [27] Z. Zhang, A. Lipton, M. Li, and A. J. Smola, Dive into Deep Learning, Free Online Resource, 2021. In: <https://d2l.ai/>.
- [28] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd ed., O'Reilly Media, 2019. In: <https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/>.
- [29] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. In: <https://www.nature.com/articles/nature14539>.
- [30] V. Nair and G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," Proceedings of the 27th International Conference on Machine Learning (ICML), Haifa, Israel, 2010, pp. 807–814. In: <https://icml.cc/2010/papers/432.pdf>.
- [31] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708. In: <https://ieeexplore.ieee.org/document/8099726>.
- [32] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440. In: <https://doi.org/10.1109/CVPR.2015.7298965>.
- [33] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," Proceedings of the 36th International Conference on Machine Learning (ICML), Long Beach, CA, USA, 2019, pp. 6105–6114. In: <https://proceedings.mlr.press/v97/tan19a.html>.
- [34] P. Goyal et al., "Accurate, large minibatch SGD: Training imagenet in 1 hour," arXiv preprint, arXiv:1706.02677, 2017. In: <https://arxiv.org/abs/1706.02677>.
- [35] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9. In: <https://doi.org/10.1109/CVPR.2015.7298594>.
- [36] C. Szegedy et al., "Inception-v4, Inception-ResNet and the impact of residual connections on learning," AAAI Conference on Artificial Intelligence, 2017. In: <https://arxiv.org/abs/1602.07261>.
- [37] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2980–2988. In: <https://ieeexplore.ieee.org/document/8237586>.

- [38] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception architecture for computer vision," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2818–2826. In: <https://arxiv.org/abs/1512.00567>.
- [39] W3C, "HTML5: A vocabulary and associated APIs for HTML and XHTML," In: <https://www.w3.org/TR/html5/>.
- [40] W3C, "Cascading Style Sheets (CSS) Overview," In: <https://www.w3.org/Style/CSS/>.
- [41] React Documentation, "A JavaScript library for building user interfaces," In: <https://reactjs.org/docs/getting-started.html>.
- [42] NumPy Documentation, "NumPy: The fundamental package for scientific computing with Python," In: <https://numpy.org/doc/stable/>.
- [43] W. McKinney, pandas Documentation, In: <https://pandas.pydata.org/pandas-docs/stable/>.
- [44] Seaborn Documentation, "Seaborn: Statistical data visualization," In: <https://seaborn.pydata.org/>.
- [45] Matplotlib Documentation, "Matplotlib: A comprehensive library for creating static, animated, and interactive visualizations in Python," In: <https://matplotlib.org/stable/contents.html>.
- [46] TensorFlow Documentation, "TensorFlow: An end-to-end open-source platform for machine learning," In: <https://www.tensorflow.org/learn>.
- [47] Ngrok Documentation, "Secure introspectable tunnels to localhost," In: <https://ngrok.com/>.
- [48] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific Data*, vol. 5, 2018. In: <https://www.nature.com/articles/sdata2018161.pdf>.
- [49] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 1251–1258. In: <https://arxiv.org/abs/1610.02357>.
- [50] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Morgan Kaufmann, 2011. In: <https://www.sciencedirect.com/book/9780123814791/data-mining-concepts-and-tech>.
- [51] D. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011. In: <https://arxiv.org/abs/2010.16061>.
- [52] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006. In: <https://doi.org/10.1016/j.patrec.2005.10.010>.

- [53] T. B. Fitzpatrick, "The validity and practicality of sun-reactive skin types I through VI," Archives of Dermatology, vol. 124, no. 6, pp. 869–871, 1988. In:
<https://jamanetwork.com/journals/jamadermatology/article-abstract/548864>.
- [54] American Academy of Dermatology, "ABCDEs of melanoma," AAD Resources, 2023. In: <https://www.aad.org/public/diseases/skin-cancer/find/at-risk/abcdes>.
- [55] K. Mader, "Skin Cancer MNIST: HAM10000 Dataset," Kaggle, 2018. In:
<https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000>.

APPENDIX A

This thesis focuses on creating machine learning models and developing a website aimed at predicting skin cancer. To promote accessibility and openness, the complete source code has been uploaded to GitHub and can be accessed through the link provided below.

*Link to the source code of SkinGuard website developed in this study:
(<https://github.com/Olivernguyen0106/SkinGuard>)*

The repository includes:

Model Code:

- Scripts for data preparation, data visualization.
- Scripts for data augmentation
- Code for training deep learning models to predict skin cancer.
- Configuration details and steps used for model training and evaluation.

Website Code:

- Source code for the website's interface and functionality.

APPENDIX B

This appendix highlights research papers and awards received during university studies, showcasing contributions to artificial intelligence, particularly in image classification and detection. The following subsections detail these publications and recognitions.

1. Conference

- **Pham Le Duc Thinh, Pham Ngoc Giau, Dang Nguyen Nam Anh, Tan Duy Le, An Mai, Nguyen Thanh Binh, and Nguyen Tan Viet Tuyen:** "*Enhancing Automated Ocular Disease Classification Based on Deep Learning: Optimization and Web-Based Deployment for Comprehensive Ophthalmic Diagnosis*", The Sixth International Conference on Real-time Intelligent Systems (RTIS 2024), Tien Giang University, Vietnam, Accepted (October 2024).
- **Pham Le Duc Thinh, Do Anh Kiet, Nguyen Thanh Binh, An Mai, and Nguyen Thi Minh Phuong:** "*A Web-Based Deep Learning Platform for Accurate and Accessible Ocular Disease Diagnosis*", The 2024 Conference on "Developing Human Resources in Artificial Intelligence (AI)", University of Science, Vietnam National University - Ho Chi Minh City (VNU-HCM), Accepted (December 2024).
- **Pham Le Duc Thinh, Nguyen Thanh Binh, Minh Khang Vo, Ho Duong Huynh Tan, Tuan Kiet Nguyen:** "*Efficient Image Retrieval and Activity Recognition in Vietnamese Daily Life using AI*", ATAC 2024 Conference on Applications of Technology, Automation, and Civil Engineering, Ho Chi Minh City, Vietnam. Accepted (November 2024).
- **Nguyen Hoang Anh Tu, Pham Le Duc Thinh, Nguyen Thanh Binh, and Loan T.T Nguyen:** "*Analysis and Implementation of Machine Learning to Forecast Vietnamese Students' Depression Levels*", The 15th National Conference on Selected Issues in Information and Communication Technology (VNICT 2022), Institute of Information Technology and Hanoi University of Industry, Hanoi, Vietnam. (Accepted 2022).

2. Awards

- **Most Popular Product Award:** "*Doctor AI*" – Honored at IT Hackathon 2024 - Solana Consumer Hack 16, sponsored by Solana, organized by the Faculty of IT, International University - HCMC, and Superteam Vietnam.
- **Innovation Award:** "*Doctor AI*" – Honored at IT Hackathon 2024 - Solana Consumer Hack 16, sponsored by Solana, organized by the Faculty of IT, International University - HCMC, and Superteam Vietnam.